

AFIT/GIR/DEM/92D-9

**AD-A258 786**



DTIC  
ELECTE  
DEC 17 1992  
S c D

IMPROVING THE QUALITY OF  
MACHINE TRANSLATIONS BY  
PROVIDING STYLISTIC GUIDELINES  
TO AUTHORS AND EDITORS

THESIS

Edward J. Walsh, Squadron Leader, RAAF

AFIT/GIR/DEM/92D-9

**92-31526**



Approved for public release; distribution unlimited

92 12 16 006

The views expressed in this thesis are those of the authors and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

UNCLASSIFIED//FORN DISSEM 1

Accession For		
NTIS GRA&I	<input checked="" type="checkbox"/>	
DTIC TAB	<input type="checkbox"/>	
Unannounced	<input type="checkbox"/>	
Justification		
By		
Distribution/		
Availability Codes		
Avail and/or		
Dist	Special	
A-1		

AFIT/GIR/DEM/92D-9

IMPROVING THE QUALITY OF MACHINE TRANSLATIONS BY  
PROVIDING STYLISTIC GUIDELINES TO AUTHORS AND EDITORS

THESIS

Presented to the Faculty of the School of Systems and Logistics

of the Air Force Institute of Technology

Air University

In Partial Fulfillment of the

Requirements for the Degree of

Master of Science in Information Resource Management

Edward J. Walsh, B.Sc., Grad. Dip. Mil. Av.

Squadron Leader, Royal Australian Air Force

December 1992

Approved for public release; distribution unlimited

## Preface

The concept for this research was born of the mating of my interest in leading edge computer applications with my delight in human expression. Although I was totally unaware of the existence of machine translation systems before stumbling across an article in the AFIT library, I am now pleased to consider myself an informed novice. My background is in Physics and Aeronautical Engineering, so my foray into the fringes of Computational Linguistics is perhaps somewhat presumptuous. Nevertheless, I believe that this thesis presents a useful, albeit small finding.

From the beginning of this project, I was ably guided by my primary thesis advisor, Dr. Charles R. (Dick) Fenno, who helped me mould a vague idea into a focused and manageable research project. Major Steve Teal was also helpful in one of his first incarnations as a thesis advisor. I am grateful for the advice and gentle suggestions provided by both.

I am also rather indebted to two representatives of commercial organizations. Colonel W.H.I. Moos, US Marine Corps (Ret.) of Globalink, Inc. very speedily provided me with a copy of the latest version of their GTS machine translation software at a critical point of the research. His speedy and generous action averted a major disruption to the project. Miss Anja Krammer of NCR, was also a great help in eliciting the cooperation of two professional translators to act as evaluators. Her involvement doubled the number of evaluators available to me.

I am also grateful to Mrs. Carole Carlson and Mrs. Federique Goodall, the other two evaluators who participated in the research. In addition, I would like to

thank all those AFIT faculty and staff who helped me along the way in some way or another. I would be remiss if I did not make special mention of the AFIT library staff who processed a large number of inter-library loan requests on my behalf.

Finally, I would like to thank my wife, Robyn, who not only acted as a proof reader, but who solidly supported my efforts from go to whoa. Without her patience and cooperation, I could not have successfully completed this project.

Edward (Ernie) J. Walsh

## Table of Contents

	Page
Preface . . . . .	ii
List of Figures . . . . .	vii
List of Tables . . . . .	viii
Abstract . . . . .	ix
I. Introduction . . . . .	1
Military Applications of Machine Translation . . . . .	2
Problem Statement . . . . .	2
Research Objectives . . . . .	2
Thesis Outline . . . . .	3
II. Literature Review . . . . .	4
Introduction . . . . .	4
Reasons for Machine Translation . . . . .	4
Alternatives to Machine Translation . . . . .	6
The History of Machine Translation . . . . .	7
Approaches to Machine Translation . . . . .	8
Operational Machine Translation Systems . . . . .	12
Current Research . . . . .	16
Research in Related Fields . . . . .	18
Facilitating Accurate Machine Translation . . . . .	19
Conclusion . . . . .	21
III. Methodology . . . . .	23
Scope of the Research . . . . .	23
Outline of Methodology . . . . .	24
Developing the Stylistic Guidelines . . . . .	25
Testing the Guidelines . . . . .	26
Conclusion . . . . .	34
IV. How the Guidelines Were Developed . . . . .	35
How the Guidelines Were Developed . . . . .	35
Validating the Guidelines . . . . .	35
The Final Guidelines . . . . .	36
List of Guidelines . . . . .	37

	Page
Guideline 1 . . . . .	38
Guideline 2 . . . . .	39
Guideline 3 . . . . .	40
Guideline 4 . . . . .	41
Guideline 5 . . . . .	43
Guideline 6 . . . . .	44
Guideline 7 . . . . .	45
Guideline 8 . . . . .	46
Guideline 9 . . . . .	47
 V. How the Guidelines Were Tested . . . . .	 48
How the Texts Were Selected . . . . .	48
How the Guidelines Were Applied to the Texts . . . . .	48
How Application of the Guidelines Was Assessed . . . . .	49
How the Texts Were Machine Translated . . . . .	50
How the Quality of the Translations Was Evaluated . . . . .	52
Comments Provided by Evaluators . . . . .	57
 VI. Discussion of Results . . . . .	 61
Discussion of Scores . . . . .	61
Interpretation of Sign Test Results . . . . .	62
Conclusion . . . . .	63
 VII. Conclusions and Recommendations . . . . .	 64
Research Summary . . . . .	64
Conclusion . . . . .	65
Recommendations . . . . .	65
Summation . . . . .	69
 Appendix A: Test Paragraphs . . . . .	 70
Appendix B: Guidelined Paragraphs . . . . .	72
Appendix C: Document Input to GTS . . . . .	77
Appendix D: Document Output by GTS . . . . .	81
Appendix E: Completed Evaluation Sheets . . . . .	85
Appendix F: Graphs of Frequency Distributions of Scores . . . . .	127
Appendix G: Statistix Sign Test Reports . . . . .	129

	Page
Bibliography . . . . .	131
Vita . . . . .	141



## List of Figures

Figure	Page
1. Illustration of Experimental Design . . . . .	27
2. Frequency Distribution for Intelligibility - Original Paragraphs . . . . .	127
3. Frequency Distribution for Intelligibility - Researcher Guidelined Paragraphs . . . . .	127
4. Frequency Distribution for Intelligibility - Volunteer Guidelined Paragraphs . . . . .	127
5. Frequency Distribution for Accuracy - Original Paragraphs . . . . .	128
6. Frequency Distribution for Accuracy - Researcher Guidelined Paragraphs . . . . .	128
7. Frequency Distribution for Accuracy - Volunteer Guidelined Paragraphs . . . . .	128

### List of Tables

Table	Page
1. Results of Random Paragraph Selection . . . . .	49
2. Frequency of Application of Guidelines . . . . .	50
3. Scores Awarded by Evaluators . . . . .	56
4. Frequency Distribution of Scores . . . . .	57
5. Sign Test P-Values . . . . .	57
6. Summary of Score Variations . . . . .	62

Abstract

This research sought to identify stylistic guidelines which could be used by the authors and editors of English language texts in order to improve the quality of subsequent machine translations to other languages. A literature review was conducted to provide background information and to determine what guidelines would be appropriate. Nine guidelines were developed on the basis of the review. The guidelines were then tested. They were applied to five paragraphs from a Royal Australian Air Force technical publication by the researcher and an independent volunteer. Both the resulting texts and the original paragraphs were translated into French using the Globalink Translation System and the quality of the translations obtained was evaluated by four human translators. While the translations of the paragraphs which had been edited by the volunteer were significantly better than the translations of the original texts, the translations of the paragraphs edited by the researcher were not. Use of the guidelines by authors and editors cannot be recommended on the basis of this result. However, analysis of the conduct and results of the research suggested potentially fruitful alternative approaches to the development and testing of guidelines. These approaches are presented as specific recommendations for further research.

# IMPROVING THE QUALITY OF MACHINE TRANSLATIONS BY PROVIDING STYLISTIC GUIDELINES TO AUTHORS AND EDITORS

## I. Introduction

As the quantity of international communications increases, so does the need for translations between the various languages used throughout the world (93:166; 106:16). However, an insufficient number of skilled human translators are available to meet these growing translation needs (65:58). Consequently, machine translation systems have been developed to perform some of the required translation. Machine translation is a term which is widely used to describe the use of machines as an aid in the translation of text from one natural language to another. In this context, a natural language is a language spoken or written by people (48:91) in order to communicate with each other.

The most capable modern machine translation systems can perform detailed analysis of input text, and can subsequently transform the original text into an accurate representation in a target language. However, even the most advanced systems make errors in translation because some input texts possess characteristics which render them difficult to analyze. Less advanced systems are even more prone to erroneous translation if the input text is difficult to analyze. In recognition of this difficulty, some researchers have suggested that the composition of texts should be controlled to facilitate more accurate machine translation (69:77; 89:35-36).

## Military Applications of Machine Translation

One of the earliest machine translation systems was developed for the USAF to translate from Russian to English (63:241; 12:6-1). Other military applications have followed. The existing and potential applications of machine translation within the military fall into three broad categories. The most common application is the translation of foreign source documents for intelligence purposes. The next most common application is the translation of documents such as maintenance publications into the language of an allied country. Finally, an application which is not yet widespread is the translation of operational documents within multinational forces. In an era of increasing international military cooperation, the application of machine translation within the military is likely to become more widespread as the systems become more capable. However, if machine translation systems are to become useful in operational military environments, they must be accurate.

## Problem Statement

Although it has long been recognized that the characteristics of input texts affect the quality of machine translations, a general set of stylistic guidelines which could be applied by authors and editors to produce texts which are predisposed to high quality machine translation from one language to another has not yet been developed.

## Research Objectives

The objectives of this research were to determine what set of stylistic guidelines should be applied to English language texts in order to improve the

quality of subsequent machine translations and to test the efficacy of the guidelines developed when applied to RAAF technical documents which are machine translated to French.

In support of these objectives, the following sub-objectives were pursued:

1. Identify passages from published sources which provide relevant guidance.
2. Develop a set of guidelines based on this guidance.
3. Test the efficacy of the guidelines.

The methodology which was applied to meet these objectives is described in chapter three.

### Thesis Outline

Chapter two of this thesis contains a review of the literature and provides further background on machine translation. In addition, the operational weaknesses of machine translation systems and the approaches taken to overcoming these weaknesses are examined. Chapter three provides a detailed description of the methodology which was applied, and elaborates on the scope of the research. Chapter four describes how the guidelines were developed and presents the proposed list of guidelines. Chapter five describes how the efficacy of the guidelines was tested, while chapter six discusses the results obtained. Finally, chapter seven presents the conclusions and recommendations arising from the research.

## II. Literature Review

### Introduction

This chapter reports the findings of a literature review performed to provide background information for this research project. The reasons for machine translation are discussed, other alternatives to human translation are examined, and the history of machine translation is summarized. The various approaches taken to machine translation are then examined and the major operational systems are briefly described. Research in machine translation and related fields is also summarized. Finally, techniques for facilitating accurate machine translation are discussed.

### Reasons for Machine Translation

Although some scholars believe that all modern languages derive from a single ancient African tongue (40:55), the differences which now exist prevent most people from mastering a second language (88:37). In an age of increasing population, increasing literacy, and increasing multinational activity, this fundamental human communication problem has led to "a definite and increasing need for translation and translation services on a worldwide basis" (93:166; 106:16). In Canada, where "the Languages Act 1969 requires all government documents and publications to be made available in English and French, . . . translators process up to one million words [per] day" (1:67). At the European Commission, which is "the administrative arm of the European Community . . . 1,000 translators churn out an estimated one million pages of text" every year (28:102-5) and "the European Parliament is spending more than half of its budget

on translation" (103:8-1). Unfortunately, there are not enough human translators available to meet these expanding translation needs (65:58; 106:16; 135:64).

Furthermore, the cost of human translation is high (41:RT-46; 138:31). "The annual bill for translation in the seven languages spoken by people who live in the countries that have five million or more inhabitants has been estimated at \$15 billion" (63:238). The appropriate use of machine translation systems can reduce these high translation costs (21:23; 62:10-14; 99:118; 104:152; 118:47).

Machine translation systems can also provide faster translations (119:33; 126:100; 133:203) which feature more consistent terminology (2:150; 99:116-117; 104:152; 137:145). Thus, translation lead time can be reduced (104:152), and productivity can be increased (102:126). Machine translation can also be convenient and secure (118:48), since those requiring translations can usually process the translation themselves without involving an outside party. It is particularly appropriate when a translation is required for assimilation rather than dissemination since the quality of the translation need not be so great (91:1-2; 106:3).

Finally, the development of machine translation systems can lead to advances in the related fields of linguistic theory, software technology, natural language processing and artificial intelligence (77:4).

Loll Rolling of the Commission of the European Communities ably summarizes the reasons for using machine translation. "The objective of machine translation is to produce translations, for someone who needs them and who is willing to pay, either to receive a quicker, a cheaper or a better translation than human translation can supply" (96:187).



## Alternatives to Machine Translation

Machine translation is only one method of coping with the large volume of translation requirements. Another approach is to write documents in a form which can be understood by more than one language group, eliminating the need for translation.

The development of the Esperanto language by Ludwig Lazarus Zamenhof in 1887 (15:418) was an early attempt at universal communication. Esperanto uses word bases common to many European languages, identifies a word's part of speech by its ending, is extremely regular in its conjugation of verbs, and features simplified inflections. Because of these features, Esperanto is easier to learn than natural languages. However, despite its advantages, the use of Esperanto has not yet become widespread, although some scholars continue to argue for its adoption as the international standard second language (11:71-76).

Another approach to universal communication is the use of various forms of simplified English (34:26-27; 13:217; 39:91). These subsets of the English language, which "include Controlled English, Simplified English, and International English; . . . [are intended to] make documents in English easy to read and understand" (13:217). To date, these systems have been used primarily to write maintenance publications intended for use in many different countries. The first of these simplified English systems was Caterpillar Fundamental English which was developed for the Caterpillar Tractor Company in 1971 (34:24). Although Simplified English is being widely used throughout the international aerospace industry (92:4-5), simplified English systems are not likely to be universally accepted because of their limited range of expression. While they are able to

describe maintenance procedures effectively, they are unable to cope with more complex language such as the description of machinery operation and design.

Universal communication through the use of a standard international language or language subset is feasible, particularly for limited applications where the intended reader can receive at least limited training. However, widespread use of such approaches is unlikely in the near future. Consequently, it will continue to be necessary to translate documents from one language to another.

### The History of Machine Translation

#### Early Efforts

Research in machine translation commenced in 1946 (116:91; 111:854). Although the research was limited by the poor performance and high cost of contemporary computers, sufficient progress was made to attract US government grants in 1954 at Georgetown University, and in 1958 at the University of Texas (63:241). "By 1965 these two systems, and another developed . . . for the US Air Force were in full operation" (63:241). However, the input overheads, post editing requirements, and high cost of computer time resulted in high cost systems (63:241). Despite the steady progress in the development of systems, the Automatic Language Processing Committee (ALPAC) of the National Research Council concluded in 1966 that machine translation was unproductive and unnecessary (63:241; 82:33; 106:1). This critical report resulted in a marked reduction in machine translation research. "By 1973, . . . only three government-funded projects were left in the US, and by late 1975 there were none" (106:1).

### The First Practical Systems

Despite the ALPAC report, some research continued, particularly in the private sector, where companies recognized a ready market (106:1). Systran, which was adopted by the USAF in 1970 for translation of Russian texts to English was among the first systems to see operational use (69:76).

### Recent Developments

During the last ten years, many other systems have been developed, and many more are currently being developed (106:1). One of the most ambitious of the newer systems is Eurotra, which is being developed by the European Community to translate between any combination of its member nations' nine languages (28:102-5). The project commenced in 1983 (7:979), and has taken 160 researchers (129:160) and approximately \$30 million (28:102-5) to develop thus far. At the other end of the cost spectrum, simple translation systems are now available for personal computers for \$2,500 or less (107:147; 98:302; 19:51).

Research has not been confined to translation of the European languages. Many systems are either operational or under development in Japan, some under the sponsorship of the Japanese government (24:102-13-102-14; 63:242). A few systems have been developed to translate between Chinese and English (51:95), and a pilot system for translation between two of the 24 Indian languages has been completed (94:522).

### Approaches to Machine Translation

Machine translation is a blanket term which includes any use of machines during the translation process. Within this broad field, a range of terminology

refers to a variety of ways of using machine translation systems (14:RT-58; 100:93; 106:2; 118:47; 117:106-107).

Machine translation systems can use either of the direct, transfer, or interlingual linguistic approaches (25:2-4; 84:23). The direct approach, which was the earliest technique developed, involves mapping the input language directly to the target language, and provides a word for word translation (25:2). Later systems used the transfer approach, which is a three phase technique involving analysis of the target language, followed by transfer to a target language symbolic representation, and finally, generation of the target language (25:2-3). Some other systems use an interlingual approach, also known as the pivot approach (42:8; 72:113), whereby a universal intermediate symbolic representation is used for all language pairs (25:3-4; 32:22-23). This eliminates the need for transfer modules for each language pair (25:3). However, it is the most difficult approach to implement (100:94).

### Levels of Translation

Irrespective of the terminology used to describe systems, or the linguistic approach used, all systems can be categorized functionally in terms of a model developed by Makoto Nagao which identifies four levels of translation (76:993-994).

Level 1. Free and creative translation which aims at the same mental reactions by the readers of the source and target languages.

Level 2. Sentence-by-sentence translation in the free sentential style. Language particularities are fully considered in translation.

Level 3. Literal translation. The sentential structures of the source language remain strongly in the target language. However, the selection of translation words is correct from the standpoint of semantics.

Level 4. Mechanical translation. The sentential structure of the target language is a crude mapping of the one in the source language, and the selection of translation word is almost one to one. (76:993-994)

These classifications are particularly useful when analyzing approaches to machine translation.

#### Level Four Translation

The earliest machine translation systems and current low priced systems conform to Nagao's fourth level. The simplest of these systems are little more than automated dictionaries which provide the operator with a choice a possible translations word by word (107:147; 10:129; 98:303-304). These systems require considerable operator intervention or post editing (98:302), but they may still be effective for language pairs "with similar structures, like Japanese and Korean, or French and Italian. But this translation level is quite unsatisfactory between languages of quite different characteristics such as Japanese and English" (76:994).

Level four systems generally do not provide natural sentence structure in the target language, and also have problems translating homonyms because they are unable to determine which foreign translation to choose. This problem also occurs in languages other than English. An example is the French word "avions" which can mean either "had" or "aircraft" (65:60).

#### Level Three Translation

Level three systems are a logical development of level four systems. The main difference is that level three systems perform some analysis of the sentence structure to determine the syntax of the sentence. This analysis, termed "parsing", yields a symbolic representation of the sentence which then aids in word selection and target language sentence structuring (10:127-128). Although these systems

are an improvement over the level four systems, they are still prone to mis-translation of sentences because they do not check the semantic validity of the translation against the original sentence. For example, one system translated the English sentence "'Daddy is hairy, like a monkey.' [as] the Chinese rejoinder 'Our long-tressed sire is fond of an ape'" (51:95). Similarly, "when [another] system translated the English sentence 'The spirit is willing but the flesh is weak' into Russian and then back to English, the resultant sentence was 'The wine is agreeable, but the meat has spoiled.'" (94:522)

### Level Two Translation

Level two translation systems attempt to form a semantic representation of sentences (43:39-44) in addition to the syntactic representation of the level three systems. A number of different approaches have been taken in achieving the overall representation, but each system forms some sort of schema representing the semantic and syntactic content of each sentence (94:521; 7:980; 10:128; 76:994; 80:73; 74:420; 33:170). Schema frameworks can be designed to suit a specific language pair or to be language independent. The advantage of language independent representations is that the same representation may then be transformed into any target language, whereas language specific representations are only suitable for unidirectional translation of a unique language pair. The systems which use these two forms of representation are known as interlingual systems and transfer systems respectively (94:521-5; 7:980; 128:24; 24:102-14; 51:95; 98:303; 79:126-138).

Although level two systems produce more accurate translations than lower level systems, some potential for mis-translation remains. The major sources of

such errors are semantically ambiguous sentences such as "I saw a man on the hill with the telescope" (94:523), or "it is advisable to avoid flying aeroplanes" (86:95), which can be interpreted in a variety of ways. Each potential meaning may have a different translation in the target language (94:523). Although a human translator can probably select the correct interpretation from the surrounding text, machine translation systems are not yet capable of such abstract inferences.

Because of the complexity of the required software, many level two systems are written in high level programming languages such as Lisp or C-Prolog (129:165). This approach leads to efficient programming and easily maintained systems. However, researchers are discovering that the inherent operational inefficiencies of the high level languages are resulting in unacceptable translation speeds for the more complex systems. Consequently, systems are being re-written in lower level languages such as C, in order to achieve an acceptable translation speed (80:82).

### Level One Translations

Machine translation systems which provide "free and creative translation which aims at the same mental reactions by the readers of the source and target languages" have not yet been developed (63:243). Consequently, level one translation remains the sole preserve of the experienced human translator for the time being.

### Operational Machine Translation Systems

Machine translation systems are currently being used throughout the commercial, public and private sectors. Users range from large organizations such as NATO and IBM to private individuals (91:1-2). The demand for operational

systems has generated considerable development effort and over a dozen machine translation systems are now available on the US market alone (23:56). Brief descriptions of some of the current systems follow.

### GTS

The Globalink Translation System (GTS) was used in the course of this research project to machine translate test paragraphs. It is a personal computer based system which can operate in either batch or interactive mode. Specialist microdictionaries are available and users can construct their own dictionaries to cater to their specific requirements. A variety of word processors can be used with the system, and it accepts input and generates output as either ASCII text or WordPerfect files. The system is also marketed with an integrated workstation featuring optical character recognition and laser printer and as a palmtop system. It is currently available for English to and from Spanish, German, and French, and for Russian to English (37; 107:147). A Chinese dictionary system is also available and an Arabic system is under development (70). One independent reviewer believes that GTS is the best personal computer based English to and from Spanish translation systems available (68:32).

### ALPS

The Automated Language Processing System (APLS) is an interactive translation system which is integrated with a word processor and was developed in the US (9:193-194). It is available for both IBM personal computers and Unix platforms (14:RT-58-RT-59) and can cater to translation from English to five different languages (90:99). It has achieved sales in Europe (97:164) as well as North America (14:RT58).



### Logos

Logos Corporation has developed systems for unusual languages such as Vietnamese (23:56) and Farsi. However, its commercial product, which operates on personal computers, concentrates upon the more traditional European languages, the first of which was German. It is a batch system which is integrated with a word processing system and claims some artificial intelligence (38:133-134).

### METAL

METAL is a high capacity transfer system which translates from German to English. It runs on a LISP machine with several translator workstations attached and is suitable for use within organizations with a high translation throughput (102:122; 103:8-2-8-4; 104:152; 115:92). The system is being extended to include translation of other romance languages (102:127).

### Systran

Systran was one of the earliest operational systems, and is now the most widely used (99:116). It was developed by LATSEC, Inc. (3:4; 36:4; 99:116) for the USAF, who started using the system in 1970 to translate Russian texts to English (10:129; 69:76). Although the original system provided little more than basic word for word translation (98:303), considerable refinements have since been made (98:303), and it now "translates through analysis of the sentence's syntax" (10:129).

Early development work was performed by two AFIT students who demonstrated that Systran could be adapted for use as a French to English translation system (3:56-57; 36:35-38). After further development, Systran was selected in 1976 by the European Community for translation between English, French, German, and Italian (28:102-5). It is now in use with large organizations

such as NASA, Xerox Corporation, and General Motors of Canada, and is available to smaller users "through on-line terminals and service bureaus" (99:116) and by modem connection to Systran Corporation (10:129; 81:20).

Systran operates on mainframe computers. It supports 15 language pairs and more are under development (99:116).

### Other Systems

In Japan, Fujitsu's ATLAS was the first commercial English to Japanese translation system and ATLAS II has since been developed to translate in both directions and with other languages (44:19; 125:317; 126:93-100). Toshiba's TRANSAC (5:107), NEC's PIVOT (71:113), Hitachi's HICATS/JE (50:101; 87:141), Mitsubishi's MELTRAN (22:217), PENSEE (22:219) and TAURUS (115:91) are among about 10 other systems currently being used in Japan to translate between English and Japanese (57:111; 78:183).

In North America, Weidner was an early US entrant to the field and is now in use for translating between a variety of languages (106:8-9). The Pan American Health Organization developed Spanam and Engspan, which has been operational since 1980, for translation between English and Spanish (130:115). TAUM-METEO has been in use even longer in Canada for translating weather forecasts (46:18). A US system called Computran has even ventured into the competitive English to Japanese market (59:82-83).

Elsewhere, the Israeli's have developed TOVNA for translation between English and French (91:1-1), and European systems include Susy, GETA (90:100), and Titus (97:164; 112:185).

Finally, a number of inexpensive PC based systems such as PC-Translator, Spanish Assistant, Toltran, Translator, Translate, and Transactive, are available (68:34; 19:51).

### Current Research

Developmental research in machine translation, which combines the "fields of linguistics, artificial intelligence, and computer programming" (42:10), continues unabated. Increasing emphasis is being placed on applying artificial intelligence and neural network methods to machine translation in an effort to translate abstract concepts (24:102-14). Efforts are also being made to apply existing natural language understanding systems to the task of machine translation (61:157).

### US Research

US researchers are not funded as well as their Japanese counterparts. However, some significant research continues. A 30 member team is working at Carnegie-Mellon University on an ambitious universal knowledge based machine translation (KBMT) system (47:36; 120:72-73). MIT researchers are also developing a universal parser based upon the theories of Noam Chomsky (40:51-52). In addition, IBM is working on basic research (21:24-25; 91:1-4). Meanwhile, the US Department of Commerce is encouraging research in translation between English and Japanese (58:177).

### European Research

The largest European project is the European Community's Eurotra. This project is aiming for translation between any combination of its member nations' nine languages (28:102-5). Although the project has been underway for almost 10

years, and has consumed considerable resources (28:102-5), it is still some way from producing an operational system (103:8-1). In the meantime, European companies are developing their own systems such as Philips' Rosetta and the Netherlands company BSO's DLT (103:8-2; 91:8-2). Universities such as the UK's Manchester Institute of Science and Technology, France's University of Grenoble, and Germany's Saarbruecken University are also engaged in research (47:35; 131:28).

### Japanese Research

At a 1989 machine translation seminar, Akira Kubota of the Ministry of International Trade and Industry stated that, "in Japan, machine translation has been of great interest. This is because we deeply feel that the Japanese language is a barrier to establishing communication with other countries" (56:170). In recognition of the communication barrier, Japan has invested heavily in machine translation research as part of the Fifth Generation Computer Project (31:13). English is not the only language being targeted. The Research Cooperation on Machine Translation Systems with Neighboring Countries project is also underway to develop systems to translate between Japanese and other Asian languages such as Chinese, Thai, Malay, and Bahasa Indonesian (47:35; 124:50-51). The Japanese division of IBM is also involved in Japanese research efforts and has developed a system which they are using in-house to translate IBM manuals (21:25).

### Research in Other Countries

Canada has a natural interest in machine translation because of its bilingual policies. Therefore, it is not surprising that the Canadian government is encouraging research (123:163). Most of the remaining research is being conducted

in Asia. China, Thailand, Malaysia, and Indonesia are all participating in the Japanese cooperative research project to develop systems to translate between their languages and Japanese (56:171). In addition, Korea is working on systems to translate from Korean to Japanese and English (53:172), and Thailand, Indonesia, and Malaysia are all working on systems to translate to and from English (108:175; 113:168; 134:113). India is also actively researching machine translation (94:522).

### Research in Related Fields

Until recently, the only input format accepted by most machine translation systems has been text files stored on computer media (10:126). However, research into optical character reader technology is now beginning to produce units which are accurate enough to channel input directly to translation systems (96:189; 101:5-2). Even systems which recognize Japanese characters have been developed (58:177). These OCR systems will further boost system productivity (128:25).

A simple system for the translation of speech has been developed by British Telecom (27:540; 20:1071; 136:1100). Other researchers are hoping to extend this work to produce a real time speech translation system in the near future (24:102-14; 49:70; 86:96; 97:164; 106:11; 135:54).

Natural language processing (NLP) occurs when a computer "takes input in human language and extracts something meaningful to a computer" (26:25). Consequently, advances in natural language processing systems such as database interfaces (26:25) might contribute to improvements in machine translation. Similarly, advances in artificial intelligence can enhance machine translation (42:9; 60:305).

## Facilitating Accurate Machine Translation

One of the greatest obstacles to accurate machine translation is poorly written input texts. The familiar rule, "garbage in, garbage out" (93:167; 103:8-11) applies. Another significant obstacle is ambiguity, one form of which is polysemy, wherein "a word in the source language has multiple meanings translating each into different words in the target language" (16:30). This includes the familiar case of homography wherein a word has "multiple syntactic categories" (16:30; 18:RT-2). For example "run" can be a noun, a verb, or an adjective (21:24). An example of non-homographic polysemy is the word "hit", which can translate into a different word in languages such as Spanish depending on what is being hit and how (16:30). This is also known as lexical ambiguity because the ambiguity hinges on the ambiguity of a particular word (21:24). Structural ambiguity is another form of ambiguity wherein the ambiguity is introduced by the syntax of a passage (21:24). For example, the statement "I am looking at the person with a telescope" (21:24) is structurally ambiguous because it can be interpreted in a number of ways.

Some systems do not even notice the ambiguity and default to one of the possible meanings. However, those systems which are capable of recognizing ambiguities tend to adopt one of two approaches. The first of these is disambiguation by asking, wherein the operator of an interactive translation system is asked to resolve the ambiguity (121:39; 138:33). This is the approach taken by the ALPS system (9:195-196). Batch systems cannot ask questions, so they disambiguate through semantic analysis (64:168-170). Other systems use a combination of both approaches to minimize time consuming interaction with the operator (85:57).

Some systems also require that the input conform to some subset of the input language (17:238-239; 54:156-161; 112:186), such as Xerox's Multinational Customized English (42:7-8; 106:7). This subset is usually known as a sublanguage, a restricted language (42:7), or a controlled language (84:37). The use of a sublanguage limits the range of structures which the system might be required to analyze, and so reduces ambiguity (54:159-161).

When input text is required to be unconstrained natural language, some other approaches can be used. Some systems build in fault tolerance, or robustness, so that the system continues to translate even when a possible translation cannot be found for a particular passage (8:472). Alternatively, input texts can be edited in accordance with some criteria after they are written. Tools such as readability formulas (29:RT-117; 52:38) and grammar checkers can be useful in this case. However, there is some doubt about how useful readability formulas are as an aid to increasing comprehensibility of text by humans (95:50). Consequently, their usefulness in conjunction with machine translation should not be assumed.

Some machine translation systems include a pre-processing module which detects portions of the text which are difficult to translate and offers the opportunity to modify the text prior to translation (6:123). These pre-processing modules are system specific and are used by machine translation system operators, rather than authors. Since operators may not be familiar with the subject matter of the text, there is some chance that the meaning of the text could be inadvertently altered during pre-processing.

In recognition of the need for a generic preprocessing tool which authors can use, IBM have developed an "Interactive Author Assistance Tool" (45:5284). IBM claims that this tool "provides an online system allowing authors of documents, which are written in one natural language, such as English, to make such documents more translatable into other natural languages" (45:5284). The system processes documents through three phases. Documents are first checked for correct spelling and are then analyzed for readability. This phase can be repeated until a satisfactory result is achieved. The document is then parsed and linguistically analyzed. Ungrammatical or ambiguous passages are highlighted for correction during this phase. Finally, the document is again checked for correct spelling (45:5285).

A final approach which can be taken if the texts must be written in natural language, and a suitable automated tool is not available, is to write in a way which maximizes translation quality without resorting to a restricted language (87:142; 93:167; 69:77; 89:35-36; 39:91-92; 72:117; 73:78). This compromise approach maintains, and perhaps even improves, the quality of the source language text while increasing the probability of a suitable translation outcome.

### Conclusion

Machine translation has progressed considerably since its inception in 1946 and future advances in computer technology are expected to lead to further significant improvements in machine translation (105:93). Some researchers believe that machine translation systems will eventually be superior to human translators in some respects (127:100). However, current systems are capable of



only adequate translation of moderately complex text, and some post editing by a human translator is still required to ensure highly accurate translations (6:123).

Accurate machine translation can be facilitated in a number of ways. One way of improving the chance of obtaining a good quality machine translation is writing texts in a way which is conducive to accurate machine translation. The remainder of this thesis reports on research undertaken to identify guidelines which authors and editors might follow in order to produce such texts.

### III. Methodology

The objectives of this research were to determine what set of stylistic guidelines should be applied to English language texts in order to improve the quality of subsequent machine translations and to test the efficacy of the guidelines developed when applied to RAAF technical documents which are machine translated to French. In order to realize these objectives, some method of developing the guidelines and subsequently testing their efficacy was required. This chapter describes the methodology which was selected.

#### Scope of the Research

The guidelines were developed on the basis of general guidance found in the literature. Therefore, they are not tailored to a particular text type, machine translation system, or target language. However, in order to ensure universal applicability of the guidelines, this research would have had to test the guidelines with the entire range of input text types, language pairs, and machine translation systems. This would clearly have been an impracticably large undertaking. In recognition of the finite time available to complete this research, the scope was limited. Specifically, this research tested the guidelines using paragraphs from a specific English language RAAF technical document, and a specific English to French machine translation system.

The RAAF technical publication which was used throughout this research was Defence Instruction (Air Force) AAP 7001.038-1, RAAF Aircraft Maintenance Philosophy and Policy. This publication was chosen because it was readily available to the researcher. It consists of eight chapters and three annexes. The

eight chapters contain a total of 386 paragraphs. The paragraphs are typically from 30 to 120 words long.

French was selected as the target language primarily because it is the foreign language with which the researcher is most familiar. It is widely used throughout the international community, which suggested that human translators would be readily available to assist in the translation evaluation phase of the research. French is also the source language of some of the RAAF's technical documentation for French sourced equipment.

Globalink Inc.'s GTS machine translation system, which is described briefly in chapter two, was used throughout this research. Because it is a PC based system, it could be used personally by the researcher rather than by a third party. This approach eliminated the possibility of external intervention during the translation process.

Globalink supplied GTS Professional Version 3.0, the latest version at the time of the research. The Professional version is somewhat more capable than the Basic version since it supports subject-specific microdictionaries which can be selected by the user when required (37). This feature was not required during the course of the research.

### Outline of Methodology

The methodology adopted for this research consisted of two phases. The first phase consisted of the development of the stylistic guidelines. This was achieved by reviewing about two hundred papers and articles for possible guidance, and then developing guidelines on the basis of the guidance found. The efficacy of

the guidelines was tested during the second phase. The individual steps which made up each phase are described in the remainder of this chapter.

### Developing the Stylistic Guidelines

The guidelines are intended for use by authors while they are composing texts which may subsequently be machine translated and by editors while revising such texts. If the guidelines are effective they will help authors and editors to produce texts which will produce higher quality machine translations than would otherwise have been the case. The form chosen for the guidelines was a concise imperative sentence supported by an explanatory passage. This form allows authors and editors to refer to the complete set of guidelines and explanations whenever they wish, while providing ready access to a short summarized list of the concise guidelines as a reminder during the production of texts.

### Identifying the Guidance

The guidance was identified by reviewing the literature to determine what guidelines had been proposed by other researchers, what characteristics of input texts lead to good machine translations, and what characteristics of input texts lead to poor machine translations.

### Developing the Guidelines

The wording of the explicitly stated guidelines was modified where necessary to accord with the concise imperative form adopted for this research. However, care was taken to ensure that the meaning of the original guidance was retained.

Wherever sources cited characteristics of texts which could lead to poor machine translations, guidelines which could eliminate those characteristics were

written by the researcher. Similarly, where sources cited characteristics of texts which could lead to good machine translations, guidelines which encourage those characteristics were written. Some degree of subjective judgement was applied in writing these guidelines. However, care was taken to ensure that the implied guidelines were firmly based upon the literature, rather than upon unsubstantiated inferences generated by the researcher.

### Ordering the Guidelines

The guidelines were derived from a number of different sources which did not suggest any objective order. Although there was no fundamental reason why the guidelines had to be listed in anything other than random order, the researcher elected to list the guidelines in what he believed was the order of difficulty of application. Consequently, the guideline which the researcher believed was easiest to apply was listed first, and the guideline which the researcher believed was the most difficult to apply was listed last.

### Validating the Guidelines

Since there was an element of subjectivity involved in deriving the guidelines from the literature, the guidelines, together with the relevant sources, were face validated by five AFIT faculty. The guidelines were then reviewed, reworded, and re-ordered in accordance with the validators' comments.

### Testing the Guidelines

Some means of testing the efficacy of the guidelines was required. This was achieved by comparing the quality of three sets of machine translations. The three sets consisted of the unaltered English text translated into French, and two sets of text altered by application of the guidelines by two separate English speaking

writers. Texts which were modified to comply with the guidelines are referred to as "guided texts" or "guided paragraphs" throughout this thesis.

#### Description of the Experimental Design

The experiment was designed to generate and compare three sets of machine translated texts which were equivalent in all respects except for variations arising from application of the guidelines. This was achieved by having two English speakers each apply the guidelines independently to a set of five English paragraphs. The original set of five paragraphs and the two sets of five guided paragraphs were then machine translated. Finally, the quality of the three sets of translations was evaluated and compared to determine whether application of the guidelines had provided any significant improvement in translation quality. Figure 1 illustrates this experimental design.

Original Texts	Application of Guidelines	Machine Translation	Evaluation of Quality
1		X	O(1)
2	X	X	O(2)
3	X	X	O(3)

Fig. 1. Illustration of Experimental Design

#### Selecting the Texts

The set of original paragraphs was formed by randomly selecting five paragraphs of 50 or more words from the RAAF publication. A sample size of five was chosen as a compromise between test accuracy and translator workload. Since

most of the translators who participated in this research were doing so voluntarily and gratis, it was considered unreasonable to expect them to dedicate large amounts of time to the evaluation of a large number of texts. An arbitrary minimum word count of 50 was selected by the researcher in the hope that this size paragraph would provide ample opportunity for application of the guidelines.

In order to randomly select the texts, sequential numbers from 1 to 386 were assigned to each paragraph in DI(AF) AAP 7001.038-1. Sub-paragraphs and lower levels of paragraph were not numbered. A random number table (67:1161) was then consulted, and the first three digits of the first random number were read. If the three digits formed a number between 1 and 386, the corresponding paragraph was selected, otherwise the next random number was read. If the paragraph was less than 50 words long, including dependent sub-paragraphs, the paragraph was rejected, and another random number was read. This process was repeated until five paragraphs of over 50 words each had been found.

#### Applying the Guidelines

The guidelines were applied to the set of five original texts by the researcher and an independent RAAF AFIT student volunteer. The volunteer was supplied with a copy of the guidelines complete with explanations and copies of the original paragraphs. He was not instructed to apply the guidelines in any particular way, but was free to interpret the guidelines as he saw fit. His application of the guidelines was intended to simulate the use of the guidelines by an average RAAF author or editor.

The researcher applied the guidelines by flagging violations of the guidelines and then correcting the violations in the order in which they appeared

in the text. Throughout the process, care was taken to ensure that the meaning of the text was preserved. Furthermore, no changes were made other than those necessary to ensure compliance with the guidelines.

#### Assessing the Application of the Guidelines

Once the guidelines had been applied to the texts, the guidelineed texts were assessed by both research advisors to ensure that the changes made were consistent with the guidelines. They also ensured that the original meanings of the texts had been retained and that no unnecessary changes had been made. However, they did not suggest any further changes to the texts since the object of the experiment was to determine what changes in translation quality could be generated by application of the guidelines by individual authors and editors.

#### Translating the Texts

After application of the guidelines by the researcher and the volunteer, there were three sets of five paragraphs each. These were the original set of five paragraphs, the set of five paragraphs guidelineed by the researcher, and the set of five paragraphs guidelineed by the volunteer. All fifteen of these paragraphs were translated into French using the GTS machine translation system. Although this system permits some forms of human assistance during interactive translation, the test texts were translated in batch mode without human intervention. The system also permits users to construct their own supplemental dictionaries in order to more accurately translate domain specific terminology. However, this facility was not used.

GTS accepts inputs and generates outputs as either WordPerfect or ASCII text files. Since the guidelines had been applied using WordPerfect 5.1, the texts



were input to GTS as a WordPerfect file in order to retain the format and underlining which were present in the English texts. The output WordPerfect file consisted of fifteen French paragraphs corresponding to the fifteen English paragraphs in the input file. French accented characters were included in the translations.

### Selecting an Instrument to Measure Translation Quality

In order to determine whether application of the guidelines had any effect on the quality of the subsequent translations, some measure of the quality of the translations was necessary. Although translation quality criteria are by nature somewhat subjective (96:190; 106:14), three criteria have been proposed (96:190). These are intelligibility, accuracy, and revision rate (96:190). Intelligibility and accuracy can be evaluated by reviewing the translations (76:1005-1007; 75:103-104). However, measurement of revision rate requires evaluators to revise each translation and measure the time taken to achieve an acceptable standard of translation (96:190; 131:34). Measurement of revision rate would have required too much of this study's volunteer evaluators' time.

The instrument selected to measure the quality of the translations was developed by Makato Nagao (76:1005-1007; 75:103-104). Nagao states that the instrument is designed to measure translation quality in terms of intelligibility and accuracy. These variables are defined as follows (76:1005-1007; 75:103-104):

**Intelligibility:** An evaluation of the extent to which the translated text can be understood by a native speaker of the target language. In Japanese to English translation, we evaluate the extent to which an average British or American reader can understand the output without any reference made to the Japanese original.

**Accuracy:** The degree to which the translated text conveys the meaning of the original text as well as a measure of the amount of difference between

the input and output sentences are evaluated. The evaluation is done by Japanese translators specializing in Japanese to English translation.

The variables are measured on scales from one to five and zero to six respectively. Nagao has provided a description of the features of text which correspond to each scale value (76:1005-1007; 75:103-104).

Nagao does not discuss his assessment of the validity of the scales, or offer an opinion on whether the scales are nominal, ordinal, interval, or ratio. His only use of the measurements is in determining percentiles of texts which achieve particular scores (76:1005-1007; 75:103-104). This is a nominal treatment. However, since the scales clearly provide a rank or order, they were treated as ordinal throughout this research. Although it is possible that the scales are interval, this could not be assumed without further information about how the instrument was validated and tested. Nagao states that

Just as there are no clear and objective criteria for evaluating the quality of Japanese to English translations done by humans, standard criteria for judging the results of machine translations have yet to be established. The evaluation methods proposed here are still in the trial stage, and much more refining and improving is still needed. (76:1007)

Despite Nagao's reservations, his is the best instrument available for objectively ranking translation quality, and treatment of the scales as ordinal rather than nominal or interval provided sufficient information to test the guidelines without compromising the validity of the conclusions of the test. Although Nagao applied the instrument to Japanese to English translations, the researcher considered that it was also suitable for the evaluation of translations between English and French.

### Evaluating the Quality of Translations

The quality of the translations was evaluated by four human translators, each of whom was provided with an evaluation package. The evaluation package included a set of instructions, a description of the measurement instrument, and fifteen evaluation sheets. Each evaluation sheet included an unmodified English paragraph, one of the three corresponding French translations, scales for intelligibility and accuracy, and a comments area.

The evaluators were asked to grade each translation in accordance with the two scales. These were blind evaluations because none of the evaluators were told which of the French translations arose from the guided texts and which arose from the original texts. The evaluators were also asked to provide any comments which they believed might be useful.

Nagao used native English speakers to evaluate intelligibility and experienced Japanese to English translators to evaluate accuracy (76:1006). However, each of the evaluators used in this research was competent in French usage as well as being competent French to English translators. Therefore, only one group was required to score both variables for the French translations.

### Interpreting the Results

The results were evaluated to determine whether there was any significant difference between the quality of the translations derived from the original paragraphs and the quality of the translations arising from the guided paragraphs. Since the only differences between the three groups of translations were due to application of the guidelines, any variation in the measured quality of translation must have been due to either the application of the guidelines or

random variations introduced by the evaluators. There was no method available to control for random variations. However, the statistical tests performed on the data took the sample size into account when calculating the significance of variations in the quality of the translations.

Because the variables measured by the instrument were ordinal, nonparametric testing of the results was appropriate. The particular test chosen was the Sign Test (110:109-110), which is used to evaluate paired ordinal data sets. The test determines how many of the values in one data set are greater than or less than the corresponding values in the other data set and then calculates the probability that any difference between the two data sets might be a result of random variations. Low probability values indicate that there is a strong possibility of a non-random difference between the data sets.

The six data sets which were compared with the Sign Test were:

1. I(O), which consisted of the intelligibility scores awarded by the four evaluators for each of the translations of the five original paragraphs (4 x 5 = 20 scores);
2. I(R), which consisted of the intelligibility scores awarded by the four evaluators for each of the translations of the five researcher guided paragraphs (4 x 5 = 20 scores);
3. I(V), which consisted of the intelligibility scores awarded by the four evaluators for each of the translations of the five volunteer guided paragraphs (4 x 5 = 20 scores);

4. A(O), which consisted of the accuracy scores awarded by the four evaluators for each of the translations of the five original paragraphs ( $4 \times 5 = 20$  scores);
5. A(R), which consisted of the accuracy scores awarded by the four evaluators for each of the translations of the five researcher guided paragraphs ( $4 \times 5 = 20$  scores);
6. A(V), which consisted of the accuracy scores awarded by the four evaluators for each of the translations of the five volunteer guided paragraphs ( $4 \times 5 = 20$  scores);

I(R) and I(V) were both compared to I(O), and A(R) and A(V) were both compared to A(O), a total of four tests.

### Conclusion

This chapter has described the methodology which was applied to this research project. Chapter four describes the application of this methodology to the development of the stylistic guidelines, while chapter five describes the evaluation of the guidelines.

#### IV. How the Guidelines Were Developed

This chapter describes how the chapter three methodology was implemented to develop the guidelines and presents the final list of guidelines.

##### How the Guidelines Were Developed

Of the many sources reviewed, only one explicitly suggested guidelines (114:1-2). However, sixteen others contained passages which could be interpreted as suggesting possible guidelines. Each of these passages was summarized and similar summaries were grouped together. These groupings, together with the explicit guidelines, were then considered one by one. A single concise imperative guideline was formulated for each grouping and an explanation of each of these guidelines was composed on the basis of the supporting passages. Where one of the explicitly expressed guidelines was equivalent to a guideline derived from one of the groupings of summaries, a single guideline was formed. The one explicitly expressed guideline which was not merged with a derived guideline was edited to conform to the format and style of the newly developed guidelines. This process generated ten guidelines.

##### Validating the Guidelines

A guideline validator's package was produced to facilitate validation of the guidelines. This package consisted of a cover sheet which described the purpose of the guidelines and how they were developed, a sheet listing all the guidelines without elaboration, a sheet for each guideline listing the guideline, its explanation, and supporting sources, and copies of the sources with supporting

passages highlighted and the corresponding pages marked with color coded tabs. One of these packages was passed to each validator.

The five AFIT faculty who acted as validators were C. Fenno (PhD), D. Vaughan (PhD), Maj R. Rice (PhD), LtCol C. Arnold, and Maj S. Teal. Each one reviewed the guidelines and supporting sources and provided comments on the guidelines' fidelity to the written sources from which they were derived. The guidelines and explanatory passages were revised on the basis of their comments. The most significant changes made were the deletion of a guideline which was not adequately supported by the sources, and re-ordering of the guidelines to more accurately reflect the relative difficulty of compliance with the guidelines. Some other relatively minor changes were also made to the wording of some of the guidelines and supporting passages, and some citation errors were corrected.

### The Final Guidelines

The final nine guidelines are presented in the following pages. A list of the nine guidelines is followed by ten pages presenting the explanations and citations for each guideline.

### List of Guidelines

1. Ensure that all words are spelt correctly.
2. Ensure that the text is grammatically correct.
3. Do not use parentheses.
4. Include redundant words rather than create ellipsis.
5. Use single word alternatives instead of phrasal verbs.
6. Do not use colloquialisms, idiomatic expressions, jargon, or metaphors.
7. Wherever possible, use words which have only one single, precise meaning.
8. Wherever possible, write short, uncomplicated sentences.
9. Minimize the use of conjunction-linked phrases.



### Guideline 1

Ensure that all words are spelt correctly.

#### Explanation

Machine translation systems will not correctly recognize words which are misspelt. Consequently, spelling errors can lead to mis-translation. Care should be taken to ensure that words are spelt in accordance with the conventions expected by the machine translation system. For example, some US systems will not recognize British spellings such as colour, tyre, or analyse.

#### Sources

This guideline was developed from 62:10-10; 66:177; 91:1-3; and 96:189.

## Guideline 2

Ensure that the text is grammatically correct.

### Explanation

Most machine translation systems depend on a grammatical analysis of each sentence to determine the part of speech of each word. This information is required to find the correct translation for words within the context of the sentence. If the source text is not grammatically correct, the system may mistranslate sentences on the basis of faulty analysis. Even subtle inaccuracies which are largely transparent to a human reader can cause problems. Since punctuation is often considered during the grammatical analysis, care should also be taken to ensure that clumsy punctuation does not confound an otherwise grammatical sentence. References such as Fowler's Dictionary of Modern English Usage (30) or grammar checking software may prove useful.

### Sources

This guideline was developed from 4:34; and 96:189.

### Guideline 3

Do not use parentheses.

#### Explanation

Many machine translation systems have trouble analyzing sentences with embedded parenthetical phrases. Consequently, it is best to avoid parenthesis altogether in order to maximize the accuracy of the machine translation.

#### Sources

This guideline was developed from 4:34; 16:30; 54:164; and 122:219.

#### Guideline 4

Include redundant words rather than create ellipsis.

#### Explanation

Although ellipsis is not ungrammatical, the inclusion of the redundant words which are implied within the sentence helps machine translation systems to determine the correct sentence structure. Consequently, more accurate translation ensues.

#### Examples

In the following examples (55:RT55-RT57), the underlined word could be omitted, but is best included in order to facilitate accurate machine translation.

The man whom my sister married is an electrician.

After a process creates a resource, any process that it starts inherits the resource handles.

A thread that is blocked on a resource by a call to this function does not return until the resource has been freed.

A semaphore is an object that is manipulated by threads that are issuing semaphore function calls.

A semaphore is an object that is manipulated by threads by issuing semaphore function calls.

Semaphores enable an application to signal the completion of certain tasks, and they control . . .

Semaphores enable an application to signal the completion of certain tasks, and to control . . .

A file system allows applications to create and to manage file objects on storage devices.

Attribute. The Attribute value that determines the file objects to be searched for.

### Sources

This guideline was developed from 54:163; 55:RT55-RT57; 73:80; 114:1-2;  
and 135:53.

## Guideline 5

Use single word alternatives instead of phrasal verbs.

### Explanation

A phrasal verb is a verb-proposition combination that acts together as a verb, such as "go on" or "come back". These phrases could be translated separately by a machine translation system instead of being considered together. Therefore, it is better to use a single word alternative which is more likely to be correctly translated.

### Examples (114:2)

"Continue" in lieu of "go on".

"Return" in lieu of "come back".

"Happen" in lieu of "come about".

"Find" in lieu of "come across".

"Learn" in lieu of "find out".

### Source

This guideline was developed from 114:2.

## Guideline 6

Do not use colloquialisms, idiomatic expressions, or metaphors.

### Explanation

Although some machine translation systems are able recognize common idiomatic expressions, those expressions which are not recognized will be translated literally. This will often produce a different meaning in the target language.

### Examples (114:1)

Idiomatic:     You'd better verify your flights before your departure.

Preferred:     You should verify your flights before your departure.

Idiomatic:     We should let everyone know of our plans.

Preferred:     We should notify everyone of our plans.

Idiomatic:     We need to get going soon.

Preferred:     We need to go soon.

Idiomatic:     I have to get ready for our party this evening.

Preferred:     I need to prepare for our party this evening.

### Sources

This guideline was developed from 62:10-10; 109:RT150; 114:1; and 132:RT6.

## Guideline 7

Wherever possible, use words which have only one single, precise meaning.

### Explanation

Very simple machine translation systems perform word by word translations. Where a single English word has multiple meanings, these systems are unable to tell which meaning is intended, and so are unable to choose from the possible target language translations. More sophisticated systems determine the part of speech for each word and are therefore able to narrow the choice of possible meanings. However, some English words still have many meanings within a particular part of speech.

One word which has different meanings both within one part of speech and in different parts of speech is "file". As a noun, "file" can mean a tool for smoothing metal or a collection of documents. "File" can also act as the verb corresponding to both nouns.

Some advanced research systems attempt to infer the exact meaning of words from the semantic context within which it occurs. However, this feature is not yet available in commercial systems. Consequently, more accurate translation is likely if most words in the source text have only one possible meaning. This will often be rather difficult to achieve.

### Sources

This guideline was developed from 16:30; 62:10-10; 65:60; 83:114; 111:854; 114:1; and 132:RT5-RT7.



## Guideline 8

Wherever possible, write short, uncomplicated sentences.

### Explanation

Long complicated sentences are difficult to analyze. Consequently, mis-translation can occur. Short uncomplicated sentences are more likely to be correctly analyzed and therefore be translated correctly. It is not possible to prescribe an exact number of allowable clauses or words per sentence. Authors must decide how short each sentence can be without making the text unnaturally telegraphic.

### Example

The first of the following sentences should be broken down into the second and third sentences (114:1):

This morning we visited the museums and this afternoon we visited the botanical gardens, both of which we thoroughly enjoyed.

This morning we visited the museums and this afternoon we visited the botanical gardens.

We thoroughly enjoyed both visits.

### Sources

This guideline was developed from 62:10-10; and 114:1.

### Guideline 9

Minimize the use of conjunction-linked phrases.

#### Explanation

Many machine translation systems have trouble determining the correct structure of sentences which contain conjunction-linked phrases. If an incorrect structure is chosen, mistranslation of the sentence is likely. Although it will not always be possible to eliminate conjunction-linked phrases completely, it is often possible to break these sentences into two shorter sentences without any loss of clarity of expression.

#### Sources

This guideline was developed from 54:164; and 73:79.

## V. How the Guidelines Were Tested

Once the set of guidelines had been developed and validated, they were tested to determine whether they could facilitate improved machine translation. This chapter describes the implementation of the testing methodology described in chapter three and presents the results obtained. Chapter six discusses the significance of the results presented in this chapter.

### How the Texts Were Selected

The test paragraphs were randomly selected from the RAAF publication, DI(AF) AAP 7001.038-1, which is a technical policy publication. It consists of eight chapters, each of which is independently numbered. For example, chapter one includes paragraphs 101 to 136, and chapter five includes paragraphs 501 to 557.

Five paragraphs were selected from the publication in accordance with the procedure described in chapter three. The results of the application of the selection procedure are displayed in Table 1. Thus, the paragraphs selected for the experiment were numbers 503, 521, 612, 304, and 313. These paragraphs are presented at Appendix A.

### How the Guidelines Were Applied to the Texts

The guidelines were applied to the texts independently by the researcher and a RAAF AFIT student volunteer. The volunteer was given a copy of the guidelines and the five test paragraphs. Both he and the researcher applied the guidelines in accordance with their own interpretation of the guidelines. They also indicated the changes made to the paragraphs and the guideline corresponding to

TABLE 1  
RESULTS OF RANDOM PARAGRAPH SELECTION

RANDOM NUMBER	CORRESPONDING PARAGRAPH
104	314 - less than 50 words.
223	503 - selected.
241	521 - selected.
421	Nil - greater than 386.
375	828 - less than 50 words.
779	Nil - greater then 386.
995	Nil - greater then 386.
963	Nil - greater then 386.
895	Nil - greater then 386.
854	Nil - greater then 386.
289	612 - selected.
635	Nil - greater then 386.
094	304 - selected.
103	313 - selected.

each change. Both versions of the guidedlined paragraphs are displayed together with corresponding original paragraphs at Appendix B.

The researcher made 33 changes, two of which invoked two guidelines at once, while the volunteer made 38. Table 2 lists the number of times each guideline (G/LINE) was applied to each paragraph (PARA) by the researcher (RES) and the volunteer (VOL).

#### How Application of the Guidelines Was Assessed

Copies of both sets of guidedlined paragraphs and the original paragraphs were passed to both thesis advisors. Although both advisors felt that they would have applied the guidelines differently, neither found any clear mis-application of

TABLE 2  
FREQUENCY OF APPLICATION OF GUIDELINES

	PARA 1		PARA 2		PARA 3		PARA 4		PARA 5		TOT
G/LINE	RES	VOL	RES	VOL	RES	VOL	RES	VOL	RES	VOL	
1	1	1	-	-	-	-	-	-	-	-	2
2	2	-	3	2	3	-	3	-	2	1	16
3	-	-	-	-	-	-	-	-	-	-	0
4	-	4	2	2	-	2	-	1	1	2	14
5	-	-	-	-	-	-	-	-	-	-	0
6	-	1	-	2	5	2	-	4	1	6	21
7	-	-	2	-	1	-	1	-	1	-	5
8	1	2	-	1	2	2	3	1	-	-	12
9	-	-	-	2	-	-	1	-	-	-	3
TOTAL	<u>4</u>	<u>8</u>	<u>7</u>	<u>9</u>	<u>11</u>	<u>6</u>	<u>8</u>	<u>6</u>	<u>5</u>	<u>9</u>	73
BY PARA	12		16		17		14		14		

the guidelines. The variation in the application of the guidelines reflects the natural variation between authors and editors which could be expected in the field.

#### How the Texts Were Machine Translated

In order to machine translate the paragraphs, access to a machine translation system was necessary. Once access to a system was established, the translation process was a straightforward operation.

### How an MT System Was Accessed

Colonel W.H.I. Moos, US Marine Corps (Ret.), who is the Program Manager for Government Programs at Globalink Inc., arranged for the loan of a copy of his company's GTS software for use in the translation phase of the project. Over 9000 copies of GTS have been sold worldwide. US military customers include the Inter-American Air Forces Academy at Homestead AFB in Florida, the Office of the Secretary of Defense, and the Defense Intelligence Agency (70). It is representative of PC based translation systems and was therefore considered suitable for use in this research.

### How GTS Was Used

The system used to translate the texts was Globalink Translation System (GTS) Professional Version 3.0, English to French, French to English. Globalink also supplied the French Business and Finance Microdictionary, which contains domain specific terminology. However, this microdictionary was not used because it was not compatible with the subject matter of the test paragraphs. GTS also allows users to build their own single word and semantic unit dictionaries to facilitate accurate translation of terminology peculiar to their area of application (35:1-2). This facility was not used because it could have masked the impact of the guidelines and could have introduced inadvertent bias on the part of the researcher.

The 15 test paragraphs were translated as a single document. Since Globalink translates sentence by sentence (35:IV-32), the presence of the three different versions of each paragraph within the same document had no affect on the translation of the texts. GTS accepts inputs as either WordPerfect or ASCII

text files (35:IV-17). Since the test paragraphs were already in WordPerfect Version 5.1 format, they were input to GTS as such. This ensured that the original format and underlining were preserved. The English document displayed at Appendix C was input in batch mode to GTS, which was installed on a Zenith 80286 based Personal Computer. After 15 minutes and 35 seconds of processing, the French translations displayed at Appendix D were completed.

### How the Quality of the Translations Was Evaluated

#### The Instrument Which Was Used

Translation quality was measured using Nagao's instrument as discussed in chapter three. The instrument measures quality in terms of the two variables: intelligibility, which is rated on a scale from one to five, and accuracy, which is rated on a scale from zero to six. Nagao's description of the variables and their corresponding scale values is reproduced below (76:1005-1007).

The Intelligibility scale provides an evaluation of the extent to which the translated text can be understood by a native speaker of the target language or a translator fluent in the target language. Thus, this scale measures the extent to which an average French speaker can understand the translation without any reference to the English original. Intelligibility is measured on a scale from one to five as follows:

- 1) The meaning of the text is clear, and there are no questions needed. Grammar, word usage, and style are all appropriate, and no rewording is needed.
- 2) The meaning of the text is clear, but there are some problems in grammar, word usage, and/or style, making the overall quality less than in 1.
- 3) The basic thrust of the text is clear, but the evaluator is not sure of some parts because of grammar and word usage problems. The problems cannot be resolved by any set procedure; a French speaker would need the assistance of an English speaker to clarify the meaning of those parts in the English original.

- 4) The text contains many grammatical and word usage problems, and the evaluator can only guess at the meaning after careful study, if at all. The quickest means of obtaining an acceptable translation would be a complete retranslation of the English text because too many revisions would be needed in the machine translation.
- 5) The text cannot be understood at all. No amount of revision would produce any meaning.

The Accuracy scale provides an evaluation of the degree to which the translated text conveys the meaning of the original text as well as a measure of the amount of difference between the input and output texts. Accuracy is measured on a scale from zero to six as follows:

- 0) The content of the input text is faithfully conveyed to the output text. The translated text would be clear to a native French speaker and no rewriting is needed.
- 1) The content of the input text is faithfully conveyed to the output text, and could be clearly understood by a native French speaker, but some rewriting is needed. The text could be corrected by a native French speaker without referring to the original text. No English language assistance would be required.
- 2) The content of the text is faithfully conveyed to the output text, but some changes are needed in word order.
- 3) While the content of the input text is generally conveyed faithfully to the output text, there are some problems with things like relationships between phrases and expressions, and with tense, voice, plurals, and the position of adverbs.
- 4) The content of the input text is not adequately conveyed to the output text. Some expressions are missing, and there are problems with the relationships between clauses, phrases and clauses, or sentence elements.
- 5) The content of the text is not conveyed to the output text. Clauses and phrases are missing.
- 6) The content of the input text is not conveyed at all. The output contains improper sentences; subjects and predicates are missing. In noun phrases, the head noun is missing, or a clause or phrase acting as a verb and modifying a noun is missing.

For both scales, a low score represents a better rating than a higher score.

The scales are a discrete measure of a continuous variable. Therefore, the



evaluators were forced to select from the values available even if they believed that the actual value fell somewhere between the available scores. Furthermore, the difference between scores is not necessarily constant. For example, the difference in intelligibility between scores of two and three may be quite dissimilar to the difference between scores of three and four. Because of these characteristics, the scales can be considered to be ordinal, but not interval or ratio. That is, the scores provide information about whether one translation is better than another, but do not reveal how much better.

#### How the Instrument Was Used

The quality of the translations was evaluated by four translators who will be referred to as Evaluators One, Evaluator Two, Evaluator Three, and Evaluator Four. Evaluator One is a Dayton based freelance translator who works with French, English, and Russian. She is a French native now living in the US. Evaluator Two is an English speaking Canadian who has studied French at graduate level. She has taught French, but is not a professional translator. Evaluators Three and Four are professional translators from Berlitz Translation Services. Their participation in this research was arranged by Miss Anja Krammer, Translation Coordinator at NCR Corporation's Workstation Product Division in Clemson, South Carolina.

As stated on page 32, each evaluator was given a copy of an evaluation package which included a set of instructions, a description of the measurement instrument, and 15 evaluation sheets. Each evaluation sheet included an English paragraph, followed by a French paragraph, scales for intelligibility and accuracy, and a comments area. Three evaluation sheets were produced for each of the five

original texts. Each of the three included the original English paragraph and one of corresponding French translations. Thus, one included the translation of the original paragraph, another included the translation of the paragraph which had been guided by the researcher, and the third included the translation of the paragraph which had been guided by the volunteer.

Since the original paragraphs were the only English texts included on the sheets, the evaluators did not have access to the guided versions of the English paragraphs, but were required to assess the translations of the guided paragraphs in terms of how well they represented the original paragraphs. Copies of the evaluation sheets returned by the evaluators are included at Appendix E in order to preserve the comments made by the evaluators.

### Evaluation Data

A summary of the intelligibility and accuracy scores awarded by the evaluators is presented in Table 3. In accordance with the scales used, a low score is better than a high score. The scores in each of the six data sets were then counted by score value to produce the frequency distribution reproduced in Table 4 and displayed graphically at Appendix F.

### Nonparametric Testing

The Statistix Sign Test (110:109-110) was used four times to compare the scores awarded to the translations of the original paragraphs to the scores awarded to translations of the guided paragraphs. Intelligibility (Researcher) and Intelligibility (Volunteer) were compared to Intelligibility (Original) and Accuracy (Researcher) and Accuracy (Volunteer) were compared to Accuracy (Original). The sign test calculates the probability that differences between data sets are due to

TABLE 3  
SCORES AWARDED BY EVALUATORS

EVAL	PARA	I(O)	I(R)	I(V)	A(O)	A(R)	A(V)
1	1	4	3 +	3 +	4	4	3 +
1	2	4	4	4	6	4 +	4 ++
1	3	4	4	4	6	6	4 ++
1	4	4	4	4	6	6	6
1	5	4	4	4	6	6	4 ++
2	1	3	3	2 +	3	3	3
2	2	4	4	3 +	4	4	4
2	3	4	3 +	2 ++	3	3	3
2	4	4	4	3 +	4	4	4
2	5	5	4 +	4 +	4	4	4
3	1	4	4	3 +	4	3 +	3 +
3	2	5	5	4 +	4	4	4
3	3	4	3 +	3 +	4	3 +	3 +
3	4	4	5 -	4	3	4 -	3
3	5	4	4	4	3	3	3
4	1	4	4	4	4	4	4
4	2	5	5	5	5	6 -	5
4	3	5	4 +	4 +	6	5 +	5 +
4	4	4	5 -	4	5	6 -	4 +
4	5	5	4 +	4 +	5	5	5

EVAL - Evaluator number

PARA - Paragraph number (1=503, 2=521, 3=612, 4=304, 5=313)

+ - Score is one category better than translation of original paragraph

++ - Score is two categories better than translation of original paragraph

- - Score is one category worse than translation of original paragraph

random variations. Therefore, a low probability value (p-value) indicates that the differences between the data sets was probably due to some factor other than random variation. The test compares each member of one data set against the corresponding member in the second data set. The sign, but not the amount of any difference between the two values is used in calculating the p-value.

TABLE 4  
FREQUENCY DISTRIBUTION OF SCORES

SCORE	I(O)	I(R)	I(V)	A(O)	A(R)	A(V)
0	-	-	-	0	0	0
1	0	0	0	0	0	0
2	0	0	2	0	0	0
3	1	4	5	4	5	7
4	14	12	12	8	8	9
5	5	4	1	3	2	3
6	-	-	-	5	5	1

The four tests were performed using Statistix Version 4.0 software. Test reports are included at Appendix G, and the p-values are summarized in Table 5.

TABLE 5  
SIGN TEST P-VALUES

	RESEARCHER	VOLUNTEER
INTELLIGIBILITY	0.1445	0.0005
ACCURACY	0.5000	0.0039

The significance of the results presented in this chapter is discussed in chapter six.

#### Comments Provided by Evaluators

As well as scores, the evaluators provided many freeform comments. Their comments are preserved on the evaluation sheets and a cover sheet returned with

the evaluations of Evaluators Three and Four, copies of which are displayed at Appendix E. Also included at Appendix E is a human translation of one of the original paragraphs which was provided by Evaluator One as an example of how it should be done.

Many of the comments are more relevant to the operation of GTS than to the application of the guidelines. Furthermore, some of the comments, particularly from Evaluators Three and Four, demonstrate a misunderstanding of the purpose of the research. This misunderstanding does not impact upon the scores awarded because they are objective assessments of the translations which are not dependent upon the purpose of the research.

Three of the four evaluators also criticized the wording of the accuracy scale. Scores 4, 5, and 6 all mention missing text. However, nothing was missing in the translations evaluated. Evaluator One avoided awarding scores of 5 for accuracy because this score refers to missing clauses and phrases. However, Evaluators Three and Four ignored the reference to missing clauses and phrases. The difference in approaches taken to the accuracy scale do not compromise the validity of the results obtained because all evaluators were consistent within their own scores. All analysis of the data compares an evaluator's assessments only to other of their own scores.

Although none of the evaluators had seen the list of guidelines, some of their comments support particular guidelines. Some other comments suggest other guidelines. The most significant comments are summarized below.

### Evaluator One

Evaluator One suggests many alternative translations and highlights many mis-translations.

### Evaluator Two

Evaluator Two suggests many possible approaches to improving the machine translations.

- Avoid usage of past participle as adjective (Sheet 1).
- Include "understood" words (Sheets 1, 7 and 15). This suggestion echoes guideline four.
- Prepositions such as "about" are ambiguous (Sheet 2). This comment is related to guideline seven.
- Hyphenate terminology such as Routine-Servicing to ensure correct translation (Sheets 4 and 10).
- Translate expressions such as "to make ready" as a single unit to ensure correct translation (Sheet 10). GTS can translate what Globalink refer to as "semantic units" if they are input to the dictionary. Other advanced systems also incorporate this feature.

### Evaluator Three

Evaluator Three points out on Sheet 4 that words ending in -ing cause problems in translation.

### Evaluator Four

Evaluator Four has two main points.

- Use short, simple sentences (Sheets 1 and 4). This comment supports guideline eight.

- Use grammatically simple sentences (Sheet 1). This is an intuitively attractive extension of guideline two.

## VI. Discussion of Results

The experiment conducted to test the efficacy of the guidelines produced 80 intelligibility and accuracy scores. This chapter discusses the significance of the scores and test results.

### Discussion of Scores

An initial examination of the scores presented in Table 3 of chapter five, and displayed graphically at Appendix F, suggests that application of the guidelines tended to improve the quality (i.e., intelligibility and accuracy) of the translations. Furthermore, it would appear that the quality of the translations of the paragraphs guided by the volunteer exceeded the quality of the translations of the paragraphs guided by the researcher. In order to better illustrate this apparent trend, the variations in scores between the guided paragraph translations and the original paragraph translations are summarized in Table 6. The values listed in the summary were obtained by subtracting the number of lower scores awarded by each evaluator for each set of paragraphs from the number of higher scores awarded.

This summary demonstrates that there were no large variations in the relative scores awarded by each evaluator. However, Evaluator Four detected slightly fewer net differences between the versions of the translations than the other three evaluators. Inspection of the data presented in Table 6 suggests once again that both the intelligibility and accuracy of the translations of the paragraphs guided by the volunteer were appreciably better than that of the translations of the original paragraphs.



TABLE 6  
SUMMARY OF SCORE VARIATIONS

	INTELLIGIBILITY		ACCURACY		TOTAL
	RES	VOL	RES	VOL	
Evaluator One	+1	+1	+1	+4	+7
Evaluator Two	+2	+5	0	0	+7
Evaluator Three	0	+3	+1	+2	+6
Evaluator Four	+1	+2	-1	+2	+4
Total	+4	+11	+1	+8	

#### Interpretation of Sign Test Results

In order to better quantify the preceding qualitative observations, four Sign Tests were performed to compare the intelligibility and accuracy of the translations of guideline paragraphs to that of the translations of the original paragraphs.

The p-values obtained are listed in Table 5 of chapter five.

The p-value is the probability that the difference between the data sets is due to random variations in the data rather than some causal factor. No matter how low the p-value, there is always a finite probability that the differences between the data sets is not caused by the factor being tested. However, for very low p-values, this probability is negligible. Generally, a p-value of less than 0.10 provides some evidence of causation, a p-value of less than 0.05 provides better evidence, and a p-value of less than 0.01 provides good evidence of causation.

Thus, it can be reasonably concluded that the volunteer's application of the guidelines to the test paragraphs improved the quality of the subsequent machine translations. The same cannot be said of the researcher's application of the guidelines. Although there may have been some improvement in the intelligibility of the translations of the researcher guided paragraphs, there is no evidence that the accuracy of the translations improved.

### Conclusion

The data collected indicates that the application of the guidelines by the volunteer significantly improved the quality of the subsequent machine translations as measured by the variables intelligibility and accuracy. However, the application of the guidelines by the researcher had little or no effect.

## VII. Conclusions and Recommendations

### Research Summary

#### Problem Statement

Although it has long been recognized that the characteristics of input texts affect the quality of machine translations (69:77; 89:35-36, among others), a general set of stylistic guidelines which could be applied by authors and editors to produce texts which are predisposed to high quality machine translation from one language to another has not yet been developed.

#### Research Objectives

The objectives of this research were to determine what set of stylistic guidelines should be applied to English language texts in order to improve the quality of subsequent machine translations and to test the efficacy of the guidelines developed when applied to RAAF technical documents which are machine translated to French.

In support of these objectives, the following sub-objectives were pursued:

1. Identify passages from published sources which provide relevant guidance.
2. Develop a set of guidelines based on this guidance.
3. Test the efficacy of the guidelines.

#### Methodology

Nine guidelines were developed on the basis of the guidance identified in the literature. These guidelines were applied to five randomly selected paragraphs from DI(AF) AAP 7001.038-1 by the researcher and a volunteer. The resulting

guided paragraphs and the original paragraphs were then machine translated to French using the Globalink Translation System. Finally, the quality of the translations was evaluated by four human translators.

### Results Obtained

The quality of the translations of the paragraphs guided by the volunteer was significantly higher than the quality of the translations of the original paragraphs. However, the quality of the translations of the paragraphs guided by the researcher was not significantly different.

### Conclusion

The experimental results obtained were indeterminate and cannot be used to predict the usefulness of the set of guidelines developed.

### Recommendations

Although the experimental results were indeterminate, analysis of the research process provides some useful insights.

#### Developing the Guidelines

Although hundreds of papers and articles were reviewed, only 17 of them contained guidance from which guidelines could be developed. Sources other than the literature may provide useful guidance.

Recommendation 1. Guidance should be sought from members of the machine translation industry.

Discussion. The only source which contained guidance in the form of specific guidelines was a two page set of notes provided by a Globalink Inc. linguist (114). These notes supported five of the nine guidelines, and were the sole source

of guideline five. Since members of the industry are familiar with the practical use of machine translation systems, they will probably have clear views about what should be done to input texts to improve the quality of machine translations. However, guidance provided could be biased and would probably be system specific.

Recommendation 2. Existing style guides and sublanguage specifications (page 20) should be reviewed for guidance.

Discussion. Companies such as Xerox Corporation use style guides specifically tailored for their own requirements. Sublanguages have also been specified for use with a number of machine translation systems. These sources may provide additional useful guidance.

Recommendation 3. Guidance should be generated through human translator analysis of machine translations.

Discussion. The evaluators who participated in this research provided comments on the machine translations which could have been used as guidance. Useful results could be expected if an instrument specifically designed to elicit guidance were developed and applied.

Recommendation 4. The differences between individual applications of guidelines should be analyzed to provide guidance.

Discussion. The quality of the translations of the paragraphs guided by the volunteer was significantly higher than that of paragraphs guided by the researcher. Systematic analysis of the variations in application of guidelines versus differences in translation quality could reveal how the guidelines are best applied. Furthermore, changes made which are not directly related to existing guidelines could provide further guidance.

Recommendation 5. The differences between original texts should be analyzed to provide guidance.

Discussion. Some texts generate higher quality translations than others, even without the use of guidelines. Analysis of the characteristics of "good" and "bad" texts could provide guidance.

#### Testing the Guidelines

The experiment conducted provided inconclusive results. Examination of the experiment leads to a number of recommendations for any future testing of guidelines.

Recommendation 6. Nagao's instrument should be refined, or a new instrument developed.

Discussion. Three of the four evaluators reported problems applying the instrument. The descriptions of the scale values should be revised, the range of values offered should be reconsidered, and rearrangement of the scores to the more intuitive ascending order of values should be evaluated. Finally, validation of any new or revised scale could provide sufficient grounds to consider the scales to be interval or ratio. This would permit more powerful statistical tests to be performed.

Recommendation 7. Test texts should be selected purposefully rather than randomly to ensure that each of a set of guidelines is exercised.

Discussion. Guidelines three and five were not violated in any of the five paragraphs randomly selected for use in this research. Therefore, these guidelines were not included in the test performed. More extensive testing should include all guidelines. Consequently, some texts may have to be purposefully

selected in order to provide for application of guidelines which have not otherwise been tested.

Recommendation 8. Translation quality should continue to be evaluated by human translators.

Discussion. No problems were encountered with the evaluations provided by the human translators. Furthermore, no good alternative evaluation technique was discovered.

Recommendation 9. Guidelines should be tested in conjunction with a variety of machine translation systems.

Discussion. Even if the experiment included in this research had provided a conclusive result, it could only have been applied to use of the guidelines with the GTS machine translation system. Unless the guidelines are being tested for use with a specific system, a variety of systems spanning the range of linguistic approaches and capabilities should be used. This approach provides for wider application of the experimental results.

Recommendation 10. Guidelines should be tested for each language pair of interest to the researcher.

Discussion. Experimental results for one language pair would not necessarily apply to another language pair. Although the guidance identified in this research dealt primarily with English as the source language, a variety of target languages were addressed. However, this does not assure that the guidelines would apply equally to any target language.

Recommendation 11. More than two writers should apply the guidelines being tested.

Discussion. If more than two writers had applied the guidelines for the experiment performed in this research, a determinate result might have been achieved. However, the evaluation required would have been prohibitive in this case. A larger sample of writers would provide information about the range of effects which can be expected of the guidelines, and would reveal whether the guidelines provide a significant benefit on average.

Recommendation 12. Large scale testing of guidelines should be conducted to provide information about the value of individual guidelines.

Discussion. The small test performed for this research could not provide information about the individual worth of guidelines. If a large number of tests were conducted, a multi-factorial analysis could be performed to determine which of the guidelines were beneficial.

### Summation

This research developed and tested a set of nine stylistic guidelines intended to improve the quality of subsequent machine translations. Although use of the guidelines by authors and editors cannot be recommended on the basis of the experimental result obtained, analysis of the conduct and results of the research suggested potentially fruitful alternative approaches to the development and testing of guidelines. These alternative approaches are presented as twelve specific recommendations for future research.



## Appendix A: Test Paragraphs

### Paragraph 1

503. In general, there is no precise method for establishing 'correct' maintenance intervals, because of the inherent variability in equipment performance and lack of complete knowledge about the behaviour of the items concerned. Instead, the process involves setting an interval based on an estimate of the 'correct' value, then progressively refining the estimate by analysis of maintenance and performance data collected for the item.

### Paragraph 2

521. Review of Routine Servicing Intervals. The routine servicing intervals for an aircraft are always reviewed as a set, in conjunction with a review of the tasks performed. Thus, the review task is essentially one of determining the continuing validity and effectiveness of the tasks performed on each item included in the set of routine servicings, and of the intervals at which they are performed. In addition the requirements must be considered in conjunction with flight servicing and DLM requirements, to ensure that each task on each item is performed at a single appropriate interval.

### Paragraph 3

612. While the need to retain capability within the RAAF relates directly to support of operations in the broadest sense, the capability to be established for a particular organizational element must be more closely aligned with the need of a particular aircraft type. Maintenance capability requirements in this case relate to the need to provide a maintenance support organization which will enhance the planned operational employment of the aircraft, both during peacetime and in time of war.

### Paragraph 4

304. Flight servicings comprise those servicings performed immediately before or after a flight, or between successive flights either to make the aircraft ready for flight or to restore it to a condition to perform further flights. Flight servicings are OLM tasks and include:

- a. before flight servicings,
- b. after flight servicings, and
- c. turnaround servicings.

#### Paragraph 5

313. The TA servicing comprises elements of both the BF and AF servicings, but does not normally include all requirements of each servicing. In general a TA servicing should include:

- a. checks of safety critical items for damage or deterioration which would not have been evident to the crew during the preceding flight,
- b. a visual examination of the exterior of the aircraft for obvious damage, and
- c. replenishment as required.

## Appendix B: Guidelined Paragraphs

### Paragraph 1

Changes made to the text during application of the guidelines are indicated by square brackets surrounding underlined text. The corresponding guideline is indicated in round bracket after each change.

#### Original Paragraph

503. In general, there is no precise method for establishing 'correct' maintenance intervals, because of the inherent variability in equipment performance and lack of complete knowledge about the behaviour of the items concerned. Instead, the process involves setting an interval based on an estimate of the 'correct' value, then progressively refining the estimate by analysis of maintenance and performance data collected for the item.

#### Guidelined by Researcher

503. In general, there is no precise method for establishing 'correct' maintenance intervals (G2) because of the inherent variability in equipment performance and [the] (G2) lack of complete knowledge about the [behavior] (G1) of the items concerned. Instead, the process begins by setting an interval based on an estimate of the 'correct' value. [This estimated interval can then be progressively refined by analyzing] (G8) maintenance and performance data collected for the item.

#### Guidelined by Volunteer

503. [Typically] (G6), there is no precise method for establishing 'correct' maintenance intervals[. This difficulty occurs] (G8) because of the inherent variability in equipment performance and [the] (G4) lack of complete knowledge about the [behavior] (G1) of the items concerned. Instead, the process involves setting an interval [that is] (G4) based on an estimate of the 'correct' value[. This estimated interval is] (G8) then progressively refined by [an] (G4) analysis of [the] (G4) maintenance and performance data collected for the item.

## Paragraph 2

Changes made to the text during application of the guidelines are indicated by square brackets surrounding underlined text. The corresponding guideline is indicated in round bracket after each change.

### Original Paragraph

521. Review of Routine Servicing Intervals. The routine servicing intervals for an aircraft are always reviewed as a set, in conjunction with a review of the tasks performed. Thus, the review task is essentially one of determining the continuing validity and effectiveness of the tasks performed on each item included in the set of routine servicings, and of the intervals at which they are performed. In addition the requirements must be considered in conjunction with flight servicing and DLM requirements, to ensure that each task on each item is performed at a single appropriate interval.

### Guidelined by Researcher

521. Review of Routine Servicing Intervals. The routine servicing intervals for an aircraft are always reviewed as a set (G2) [together] (G7) with a review of the tasks performed. Thus, the review task is essentially one of determining the continuing validity and effectiveness of both (G4) the tasks performed on each item included in the set of routine servicings, and the intervals at which [the tasks] (G4) are performed. In addition~~[,]~~ (G2) the requirements must be considered [together] (G7) with flight servicing and DLM requirements (G2) to ensure that each task on each item is performed at a single appropriate interval.

### Guidelined by Volunteer

521. Review of Routine Servicing Intervals. The routine servicing intervals for an aircraft are always reviewed as a set, in conjunction with a review of the tasks [which are] (G4) performed. (G6) The review task is essentially one of determining the continuing validity and effectiveness of the tasks performed on each item included in the set of routine servicings~~].~~ The review then examines] (G9) the intervals at which [valid tasks for each item] (G9) are performed. In addition~~[,]~~ (G2) the requirements must be considered in conjunction with [the] (G4) flight servicing and [Depot Level Maintenance] (G6) requirements~~].~~ This consideration will] (G8) ensure that [all tasks for each item are] (G2) performed at a single appropriate interval.

### Paragraph 3

Changes made to the text during application of the guidelines are indicated by square brackets surrounding underlined text. The corresponding guideline is indicated in round bracket after each change.

#### Original Paragraph

612. While the need to retain capability within the RAAF relates directly to support of operations in the broadest sense, the capability to be established for a particular organizational element must be more closely aligned with the need of a particular aircraft type. Maintenance capability requirements in this case relate to the need to provide a maintenance support organization which will enhance the planned operational employment of the aircraft, both during peacetime and in time of war.

#### Guidelined by Researcher

612. [The] (G8) need to retain [maintenance capability] (G6) within the RAAF relates directly to [the] (G2) support of [aircraft operations] (G6) in the broadest sense. [However,] (G8) the [maintenance capability] (G6) to be established for a particular organizational element must be (G2) closely aligned with the [needs] (G2) of a particular aircraft type. Maintenance capability requirements in this case relate to the need to provide a maintenance support organization which will enhance the planned operational [use] (G6/G7) of the aircraft, both during peacetime and [during wartime] (G6).

#### Guidelined by Volunteer

612. (G8) The need to retain capability within the [Royal Australian Air Force] (G6) relates directly to [the] (G4) support of operations in the broadest sense[. However,] (G8) the capability to be established for a particular organizational element must be more closely aligned with the need of a particular aircraft type. [The] (G4) maintenance capability requirements in this case relate to the need to provide a maintenance support organization which will enhance the planned operational employment of the aircraft, both during [times of peace and] (G6) of war.

#### Paragraph 4

Changes made to the text during application of the guidelines are indicated by square brackets surrounding underlined text. The corresponding guideline is indicated in round bracket after each change.

#### Original Paragraph

304. Flight servicings comprise those servicings performed immediately before or after a flight, or between successive flights either to make the aircraft ready for flight or to restore it to a condition to perform further flights. Flight servicings are OLM tasks and include:

- a. before flight servicings,
- b. after flight servicings, and
- c. turnaround servicings.

#### Guidelined by Researcher

304. Flight servicings [are] (G7) those servicings performed immediately before or after a flight, or between successive flights. (G8) They are performed [to either] (G2) make the aircraft ready for flight or (G2) restore it to a [flyable] (G2) condition. Flight servicings are OLM tasks. [They] (G8/G9) include (G8) before flight servicings, after flight servicings, and turnaround servicings.

#### Guidelined by Volunteer

304. Flight servicings comprise those servicings [which are] (G4) performed immediately before or after a flight, or between successive flights. [Flight servicings] (G8) either [prepare] (G6) the aircraft (G6) for flight or restore it to a condition to perform further flights. Flight servicings are [Operational Level Maintenance] (G6) tasks and include:

- a. before flight servicings,
- b. after flight servicings, and
- c. turn around (G6) servicings.

## Paragraph 5

Changes made to the text during application of the guidelines are indicated by square brackets surrounding underlined text. The corresponding guideline is indicated in round bracket after each change.

### Original Paragraph

313. The TA servicing comprises elements of both the BF and AF servicings, but does not normally include all requirements of each servicing. In general a TA servicing should include:

- a. checks of safety critical items for damage or deterioration which would not have been evident to the crew during the preceding flight,
- b. a visual examination of the exterior of the aircraft for obvious damage, and
- c. replenishment as required.

### Guidelined by Researcher

313. The TA servicing includes (G7) elements of both the BF and AF servicings, but does not normally include all the (G2) requirements of each servicing. In general, (G2) a TA servicing should include:

checks of safety critical items for damage or deterioration which would not have been obvious (G6) to the crew during the preceding flight,

a visual examination of the exterior of the aircraft for obvious damage, and

the replenishment of fluids and gases if necessary (G4).

### Guidelined by Volunteer

313. The Turn Around (G6) servicing includes (G2) elements of both the Before Flight (G6) and After Flight (G6) servicings, but it (G4) does not normally include all of the (G4) requirements of each servicing. Typically (G6) a Turn Around (G6) servicing should include:

- a. checks of critical safety (G6) items for damage or deterioration which would not have been evident to the crew during the preceding flight,
- b. a visual examination of the exterior of the aircraft for obvious damage, and
- c. replenishment as required.

### Appendix C: Document Input to GTS

503. In general, there is no precise method for establishing 'correct' maintenance intervals, because of the inherent variability in equipment performance and lack of complete knowledge about the behaviour of the items concerned. Instead, the process involves setting an interval based on an estimate of the 'correct' value, then progressively refining the estimate by analysis of maintenance and performance data collected for the item.

521. Review of Routine Servicing Intervals. The routine servicing intervals for an aircraft are always reviewed as a set, in conjunction with a review of the tasks performed. Thus, the review task is essentially one of determining the continuing validity and effectiveness of the tasks performed on each item included in the set of routine servicings, and of the intervals at which they are performed. In addition the requirements must be considered in conjunction with flight servicing and DLM requirements, to ensure that each task on each item is performed at a single appropriate interval.

612. While the need to retain capability within the RAAF relates directly to support of operations in the broadest sense, the capability to be established for a particular organizational element must be more closely aligned with the need of a particular aircraft type. Maintenance capability requirements in this case relate to the need to provide a maintenance support organization which will enhance the planned operational employment of the aircraft, both during peacetime and in time of war.

304. Flight servicings comprise those servicings performed immediately before or after a flight, or between successive flights either to make the aircraft ready for flight or to restore it to a condition to perform further flights. Flight servicings are OLM tasks and include:

- a. before flight servicings,
- b. after flight servicings, and
- c. turnaround servicings.



313. The TA servicing comprises elements of both the BF and AF servicings, but does not normally include all requirements of each servicing. In general a TA servicing should include:

- a. checks of safety critical items for damage or deterioration which would not have been evident to the crew during the preceding flight,
- b. a visual examination of the exterior of the aircraft for obvious damage, and
- c. replenishment as required.

503. In general, there is no precise method for establishing 'correct' maintenance intervals because of the inherent variability in equipment performance and the lack of complete knowledge about the behavior of the items concerned. Instead, the process begins by setting an interval based on an estimate of the 'correct' value. This estimated interval can then be progressively refined by analyzing maintenance and performance data collected for the item.

521. Review of Routine Servicing Intervals. The routine servicing intervals for an aircraft are always reviewed as a set together with a review of the tasks performed. Thus, the review task is essentially one of determining the continuing validity and effectiveness of both the tasks performed on each item included in the set of routine servicings, and the intervals at which the tasks are performed. In addition, the requirements must be considered together with flight servicing and DLM requirements to ensure that each task on each item is performed at a single appropriate interval.

612. The need to retain maintenance capability within the RAAF relates directly to the support of aircraft operations in the broadest sense. However, the maintenance capability to be established for a particular organizational element must be closely aligned with the needs of a particular aircraft type. Maintenance capability requirements in this case relate to the need to provide a maintenance support organization which will enhance the planned operational use of the aircraft, both during peacetime and during wartime.

304. Flight servicings are those servicings performed immediately before or after a flight, or between successive flights. They are performed to either make the aircraft ready for flight or restore it to a flyable condition. Flight servicings are OLM tasks. They include before flight servicings, after flight servicings, and turnaround servicings.

313. The TA servicing includes elements of both the BF and AF servicings, but does not normally include all the requirements of each servicing. In general, a TA servicing should include:

checks of safety critical items for damage or deterioration which would not have been obvious to the crew during the preceding flight,

a visual examination of the exterior of the aircraft for obvious damage, and

the replenishment of fluids and gases if necessary.

503. Typically, there is no precise method for establishing 'correct' maintenance intervals. This difficulty occurs because of the inherent variability in equipment performance and the lack of complete knowledge about the behavior of the items concerned. Instead, the process involves setting an interval that is based on an estimate of the 'correct' value. This estimated interval is then progressively refined by an analysis of the maintenance and performance data collected for the item.

521. Review of Routine Servicing Intervals. The routine servicing intervals for an aircraft are always reviewed as a set, in conjunction with a review of the tasks which are performed. The review task is essentially one of determining the continuing validity and effectiveness of the tasks performed on each item included in the set of routine servicings. The review then examines the intervals at which valid tasks for each item are performed. In addition, the requirements must be considered in conjunction with the flight servicing and Depot Level Maintenance requirements. This consideration will ensure that all tasks for each item are performed at a single appropriate interval.

612. The need to retain capability within the Royal Australian Air Force relates directly to the support of operations in the broadest sense. However, the capability to be established for a particular organizational element must be more closely aligned with the need of a particular aircraft type. The maintenance capability requirements in this case relate to the need to provide a maintenance support organization which will enhance the planned operational employment of the aircraft, both during times of peace and of war.

304. Flight servicings comprise those servicings which are performed immediately before or after a flight, or between successive flights. Flight servicings either prepare the aircraft for flight or restore it to a condition to perform further flights. Flight servicings are Operational Level Maintenance tasks and include:

- a. before flight servicings,
- b. after flight servicings, and
- c. turn around servicings.

313. The Turn Around servicing includes elements of both the Before Flight and After Flight servicings, but it does not normally include all of the requirements of each servicing. Typically a Turn Around servicing should include:

- a. checks of critical safety items for damage or deterioration which would not have been evident to the crew during the preceding flight,
- b. a visual examination of the exterior of the aircraft for obvious damage, and
- c. replenishment as required.

#### Appendix D: Document Output by GTS

503. En général, il y a aucune méthode précise pour n'établir 'correct' intervalles d'entretien, à cause de la variabilité inhérente dans le manque et performance d'équipement de connaissance complète environ le comportement des articles concernait. Plutôt, le procédé entraîne fixer une intervalle basé sur une estimation du 'correcte' valeur, alors raffinant progressivement l'estimation par l'analyse de données de performance et entretien recueillies pour l'article.

521. La revue de Routine Révisant Des intervalles. La routine révisant des intervalles pour un avion sont toujours revus comme un ensemble, dans la conjonction avec une revue des tâches exécutés. Ainsi, la tâche de revue est essentiellement une de déterminer l'efficacité et validité continue des tâches exécutée sur chaque article incluse dans l'ensemble de révisions de routine, et des intervalles à lesquelles ils sont exécutés. Au surplus les exigences doivent être considérées dans la conjonction avec le vol révisant et DLM exigences, à assurer que chaque tâche sur chaque article est exécutée à une intervalle appropriée seule.

612. Tandis que le besoin de retenir l'aptitude dans le RAAF lie soutenir directement d'opérations dans le sens le plus large, l'aptitude à être établie pour un particulier organizational élément doit être plus alignée étroitement avec le besoin d'un type particulier d'avion. Les exigences d'aptitude d'entretien dans ce cas lient au besoin à fournir une organisation de soutien d'entretien qui améliorera l'emploi opérationnel planifié de l'avion, les deux durant peacetime et à temps de guerre.

304. Les révisions de vol comprennent ces révisions exécutaient aussitôt avant ou après un vol, ou entre des vols successifs l'un ou l'autre à faire l'avion prêt pour le vol ou à le restaurer à une condition à exécuter davantage vols. Les révisions de vol sont OLM tâches et incluent:

- a. avant vol révise,
- b. après le vol révise, et
- c. turnaround révise.

313. Le TA révisant comprend des éléments des deux le BF et AF révisions, mais n'inclut pas normalement toutes exigences de chacun révisant. En général un TA révision devrait inclure:

- a. les chèques de sécurité articles critiques pour le dommages ou détérioration que n'aurait pas été évidente à l'équipage durant le vol précédent,
- b. un examen visuel de l'extérieur de l'avion pour le dommages évident, et
- c. replenishment comme requis.

503. En général, il y a aucune méthode précise pour n'établir 'correct' intervalles d'entretien à cause de la variabilité inhérente dans la performance d'équipement et le manque de connaissance complète environ le comportement des articles concernait. Plutôt, le procédé commence en fixant une intervalle basé sur une estimation du 'correcte' valeur. Cela estimait l'intervalle peut alors être raffiné progressivement en analysant des données de performance et entretien recueillies pour l'article.

521. La revue de Routine Révisant Des intervalles. La routine révisant des intervalles pour un avion sont toujours revus comme un ensemble ensemble avec une revue des tâches exécutés. Ainsi, la tâche de revue est essentiellement une de déterminer l'efficacité et validité continue des deux les tâches exécutée sur chaque article incluse dans l'ensemble de révisions de routine, et les intervalles à lesquelles les tâches sont exécutées. Au surplus, les exigences doivent être considérées ensemble avec le vol révisant et DLM exigences à assurer que chaque tâche sur chaque article est exécutée à une intervalle appropriée seule.

612. Le besoin de retenir l'aptitude d'entretien dans le RAAF lie directement au soutien d'opérations d'avion dans le sens le plus large. Cependant, l'aptitude d'entretien à être établie pour un particulier organizational élément doit être alignée étroitement avec les besoins d'un type particulier d'avion. Les exigences d'aptitude d'entretien dans ce cas lient au besoin à fournir une organisation de soutien d'entretien qui améliorera l'emploi opérationnel planifié de l'avion, les deux durant peacetime et durant wartime.

304. Les révisions de vol sont ces révisions exécutaient aussitôt avant ou après un vol, ou entre des vols successifs. Ils sont exécutés à l'un ou l'autre faits l'avion prêt pour le vol ou restaurés l'à un flyable condition. Les révisions de vol sont OLM tâches. Ils incluent avant révisions de vol, après des révisions de vol, et turnaround révisions.

313. Le TA révisant inclut des éléments dus deux le BF et AF révisions, mais n'inclut pas normalement toutes les exigences de chacun révisant. En général, un TA révision devrait inclure:

les chèques de sécurité articles critiques pour le dommages ou détérioration que n'aurait pas été évidente à l'équipage durant le vol précédent,

un examen visuel de l'extérieur de l'avion pour le dommages évident, et

le replenishment de fluides et gazer si nécessaires.

503. Typiquement, il y a aucune méthode précise pour n'établir 'correct' intervalles d'entretien. Cette difficulté survient à cause de la variabilité inhérente dans la performance d'équipement et le le manque de connaissance complète environ le comportement des articles concernait. Plutôt, le procédé entraîne fixer une intervalle qui est basée sur une estimation du 'correcte' valeur. Cela estimait l'intervalle est alors raffiné progressivement par une analyse des données de performance et entretien recueillies pour l'article.

521. La revue de Routine Révisant Des intervalles. La routine révisant des intervalles pour un avion sont toujours revus comme un ensemble, dans la conjonction avec une revue des tâches qui sont exécutés. La tâche de revue est essentiellement une de déterminer l'efficacité et validité continue des tâches exécutée sur chaque article incluse dans l'ensemble de révisions de routine. La revue examine alors les intervalles à lesquelles tâches valables pour chaque article sont exécutés. Au surplus, les exigences doivent être considérées dans la conjonction avec la révision de vol et Niveau De dépôt exigences D'entretien. Cette considération assurera que toutes tâches pour chaque article sont exécutées à une intervalle appropriée seule.

612. Le besoin de retenir l'aptitude dans la Force Royale D'air Australien lie directement au soutien d'opérations dans le sens le plus large. Cependant, l'aptitude à être établie pour un particulier organizational élément doit être plus alignée étroitement avec le besoin d'un type particulier d'avion. Les exigences d'aptitude d'entretien dans ce cas lient au besoin à fournir une organisation de soutien d'entretien qui améliorera l'emploi opérationnel planifié de l'avion, les deux durant des temps de paix et de guerre.

304. Les révisions de vol comprennent ces révisions qui sont exécutées aussitôt avant ou après un vol, ou entre des vols successifs. Les révisions de vol l'un ou l'autre préparent l'avion pour le vol ou le restaurent à une condition à exécuter davantage vols. Les révisions de vol sont tâches Opérationnelles D'entretien De niveau et incluent:

- a. avant vol révisé,
- b. après le vol révisé, et
- c. la tournure autour des révisions.

313. La Tournure Autour réviser inclut des éléments des deux L'avant Vol et Après des révisions De vol, mais il n'inclut pas normalement toutes les exigences de chacun révisant. Typiquement une Tournure Autour la révision devrait inclure:

- a. les chèques d'articles critiques de sécurité pour le dommages ou détérioration que n'aurait pas été évidente à l'équipage durant le vol précédent,
- b. un examen visuel de l'extérieur de l'avion pour le dommages évident, et
- c. replenishment comme requis.

### Appendix E: Completed Evaluation Sheets

This Appendix presents the comments made by the evaluators. The first three pages show the covering material which was sent to each evaluator with the evaluation sheets. Of the next 38 pages, 36 are copies of the evaluation sheets which were completed by the four evaluators. They are grouped by evaluator and are ordered from Evaluator One to Evaluator Four. Evaluation sheets which did not feature comments are not included. In addition to the evaluation sheets, a sample human translation of one of the original paragraphs is included at the end of the Evaluator One section. This translation was provided by Evaluator One to demonstrate the product generated by a human translator. A cover sheet which was returned with the Evaluator Three and Four evaluations is also included just before the Evaluator Three sheets.



## EVALUATION PACKAGE FOR TRANSLATION EVALUATORS

Attached are fifteen evaluation sheets. Each sheet consists of an English text, a French machine translation of the text, and an evaluation section. There are three sheets for each English text, each with a slightly different French translation. Please consider each of the French translations in terms of the Intelligibility and Accuracy scale descriptions which follow and circle the appropriate scores for each translation on the scales in the evaluation section of each sheet. Please try to be as consistent as possible in your assessments, but do not discuss your scores with anyone else evaluating the translations. To be valid, each assessment must reflect an individual viewpoint.

Comments on each translation are welcome. Any observations about constructions within the English text which seem to have led to poor translation, or shortcomings in the machine's apparent translation approach would be particularly welcome.

Your participation in this research is appreciated. If you would like a summary of the research findings, I will be happy to forward one later this year. If you would like a copy of the summary or are willing to be contacted to discuss your assessments, please provide your details below.

Name:

Organization:

Job Title:

Contact Telephone Number/s:  
(Business and/or home)

Postal Address:



Brian Walsh  
Squadron Leader  
Royal Australian Air Force  
Ph: (513) 427 3383

## INTELLIGIBILITY

The Intelligibility scale provides an evaluation of the extent to which the translated text can be understood by a native speaker of the target language or a translator fluent in the target language. Thus, this scale measures the extent to which an average French speaker can understand the translation without any reference to the English original. Intelligibility is measured on a scale from one to five as follows:

- 1) The meaning of the text is clear, and there are no questions needed. Grammar, word usage, and style are all appropriate, and no rewording is needed.
- 2) The meaning of the text is clear, but there are some problems in grammar, word usage, and/or style, making the overall quality less than in 1.
- 3) The basic thrust of the text is clear, but the evaluator is not sure of some parts because of grammar and word usage problems. The problems cannot be resolved by any set procedure; a French speaker would need the assistance of an English speaker to clarify the meaning of those parts in the English original.
- 4) The text contains many grammatical and word usage problems, and the evaluator can only guess at the meaning after careful study, if at all. The quickest means of obtaining an acceptable translation would be a complete retranslation of the English text because too many revisions would be needed in the machine translation.
- 5) The text cannot be understood at all. No amount of revision would produce any meaning.

## ACCURACY

The Accuracy scale provides an evaluation of the degree to which the translated text conveys the meaning of the original text as well as a measure of the amount of difference between the input and output texts. Accuracy is measured on a scale from zero to six as follows:

- 0) The content of the input text is faithfully conveyed to the output text. The translated text would be clear to a native French speaker and no rewriting is needed.
- 1) The content of the input text is faithfully conveyed to the output text, and could be clearly understood by a native French speaker, but some rewriting is needed. The text could be corrected by a native French speaker without referring to the original text. No English language assistance would be required.
- 2) The content of the text is faithfully conveyed to the output text, but some changes are needed in word order.
- 3) While the content of the input text is generally conveyed faithfully to the output text, there are some problems with things like relationships between phrases and expressions, and with tense, voice, plurals, and the position of adverbs.
- 4) The content of the input text is not adequately conveyed to the output text. Some expressions are missing, and there are problems with the relationships between clauses, phrases and clauses, or sentence elements.
- 5) The content of the text is not conveyed to the output text. Clauses and phrases are missing.
- 6) The content of the input text is not conveyed at all. The output contains improper sentences; subjects and predicates are missing. In noun phrases, the head noun is missing, or a clause or phrase acting as a verb and modifying a noun is missing.

## EVALUATION SHEET 1

Once you have read and understood the preceding instructions, please circle the appropriate score for each scale in the evaluation section.

### English Text

503. In general, there is no precise method for establishing 'correct' maintenance intervals, because of the inherent variability in equipment performance and lack of complete knowledge about the behaviour of the items concerned. Instead, the process involves setting an interval based on an estimate of the 'correct' value, then progressively refining the estimate by analysis of maintenance and performance data collected for the item.

### French Translation

503. En général, il y a aucune méthode précise pour n'établir 'correct' intervalles d'entretien, à cause de la variabilité inhérente dans le manque de performance d'équipement de connaissance complète environ le comportement des articles concernait. Plutôt, le procédé entraîne fixer un intervalle basé sur une estimation d'une 'correcte' valeur, alors raffinant progressivement l'estimation par l'analyse des données de performance et d'entretien recueillies pour l'article.

### Evaluation Section

Intelligibility: 1 2 3 (4) 5

Accuracy: 0 1 2 3 (4) 5 6

Comments:

As far as accuracy is concerned, I would have used the score '6' to say that the content of the input text is not conveyed at all, but no elements are missing.

'x' refers to any type of mistake.

## EVALUATION SHEET 2

Once you have read and understood the preceding instructions, please circle the appropriate score for each scale in the evaluation section.

### English Text

503. In general, there is no precise method for establishing 'correct' maintenance intervals, because of the inherent variability in equipment performance and lack of complete knowledge about the behaviour of the items concerned. Instead, the process involves setting an interval based on an estimate of the 'correct' value, then progressively refining the estimate by analysis of maintenance and performance data collected for the item.

### French Translation

503. En général, il y a aucune méthode précise pour établir 'correct' intervalles d'entretien à cause de la variabilité inhérente dans la performance d'équipement et le manque de connaissance complète environ le comportement des articles concernait. Plutôt, le procédé commence en fixant un intervalle basé sur une estimation d'une 'correcte' valeur. Cela s'estimait l'intervalle peut alors être raffiné progressivement en analysant des données de performance et d'entretien recueillies pour l'article.

### Evaluation Section

Intelligibility: 1 2 ③ 4 5

Accuracy: 0 1 2 3 ④ 5 6

Comments:

In this text, the score '4' for accuracy is used as it is, that is to say the content is not adequately conveyed. No elements are missing.

### EVALUATION SHEET 3

Once you have read and understood the preceding instructions, please circle the appropriate score for each scale in the evaluation section.

#### English Text

503. In general, there is no precise method for establishing 'correct' maintenance intervals, because of the inherent variability in equipment performance and lack of complete knowledge about the behaviour of the items concerned. Instead, the process involves setting an interval based on an estimate of the 'correct' value, then progressively refining the estimate by analysis of maintenance and performance data collected for the item.

#### French Translation

503. Typiquement, il n'y a aucune méthode précise pour n'établir 'correct' intervalles d'entretien. Cette difficulté survient à cause de la variabilité inhérente dans la performance d'équipement et le manque de connaissance complète environ le comportement des articles concernait. Plutôt, le procédé entraîne fixer un intervalle qui est basé sur une estimation du 'correcte' valeur. Cela estimait l'intervalle est alors raffiné progressivement par une analyse des données de performance et d'entretien recueillies pour l'article.

#### Evaluation Section

Intelligibility: 1 2 ③ 4 5

Accuracy: 0 1 2 ③ 4 5 6

#### Comments:

The improvement in the accuracy score does not really mean that the text is properly translated. A French-speaking reader would have a hard time understanding the text. I used a higher score to show that there is some improvement compared to the previous versions.

#### EVALUATION SHEET 4

Once you have read and understood the preceding instructions, please circle the appropriate score for each scale in the evaluation section.

#### English Text

521. Review of Routine Servicing Intervals. The routine servicing intervals for an aircraft are always reviewed as a set, in conjunction with a review of the tasks performed. Thus, the review task is essentially one of determining the continuing validity and effectiveness of the tasks performed on each item included in the set of routine servicings, and of the intervals at which they are performed. In addition the requirements must be considered in conjunction with flight servicing and DLM requirements, to ensure that each task on each item is performed at a single appropriate interval.

#### French Translation

521. La revue de Routine Révisant Des intervalles. La routine révisant des intervalles pour un avion sont toujours revus comme un ensemble, dans la conjonction avec une revue des tâches exécutés. Ainsi, la tâche de revue est essentiellement une de déterminer l'efficacité et la validité continue des tâches exécutées sur chaque article incluse dans l'ensemble de révisions de routine et des intervalles auxquelles ils sont exécutés. Aux surplus les exigences doivent être considérées dans la conjonction avec les vols révisant et les DLM exigences, à assurer que chaque tâche sur chaque article est exécutée à un intervalle appropriée seule.

#### Evaluation Section

Intelligibility: 1 2 3 ④ 5

Accuracy: 0 1 2 3 4 5 ⑥

#### Comments:

For example, I would have translated the title into "Etudier des cadences d'entretien journalier".  
Some remarks as before. The content is not conveyed at all. However, no elements are missing.

## EVALUATION SHEET 6

Once you have read and understood the preceding instructions, please circle the appropriate score for each scale in the evaluation section.

### English Text

521. Review of Routine Servicing Intervals. The routine servicing intervals for an aircraft are always reviewed as a set, in conjunction with a review of the tasks performed. Thus, the review task is essentially one of determining the continuing validity and effectiveness of the tasks performed on each item included in the set of routine servicings, and of the intervals at which they are performed. In addition the requirements must be considered in conjunction with flight servicing and DLM requirements, to ensure that each task on each item is performed at a single appropriate interval.

### French Translation

521. La revue des routines révisant des intervalles. La routine révisant des intervalles pour un avion sont toujours revus comme un ensemble, dans la conjonction avec une revue des tâches qui sont exécutés. La tâche de revue est essentiellement une de déterminer l'efficacité et la validité continue des tâches exécutées sur chaque article incluse dans l'ensemble de révisions de routine. La revue examine alors les intervalles auxquelles les tâches valables pour chaque article sont exécutés. Aux surplus les exigences doivent être considérées dans la conjonction avec la révision de vol et Niveau de dépôt d'exigences d'entretien. Cette considération assure que toutes les tâches pour chaque article sont exécutées à un intervalle appropriée seule.

### Evaluation Section

Intelligibility: 1 2 3 (4) 5

Accuracy: 0 1 2 3 (4) 5 6

#### Comments:

Same as before.  
There is a slight improvement as far as the choice of words is concerned (for example, 'DLM requirements' needed to be expanded in order to be understood), but the translation is still poor ('DLM requirements' should have been - suggestion - translated into 'normes d'entretien en dépôt (DLM)'), and the title is still misleading.



## EVALUATION SHEET 8

Once you have read and understood the preceding instructions, please circle the appropriate score for each scale in the evaluation section.

### English Text

612. While the need to retain capability within the RAAF relates directly to support of operations in the broadest sense, the capability to be established for a particular organizational element must be more closely aligned with the need of a particular aircraft type. Maintenance capability requirements in this case relate to the need to provide a maintenance support organization which will enhance the planned operational employment of the aircraft, both during peacetime and in time of war.

### French Translation

612. Le besoin de retenir l'aptitude d'entretien dans le RAAF lie directement au soutien d'opérations d'avion dans le sens le plus large. Cependant, l'aptitude d'entretien à être établie pour un particulier organisationnel élément doit être alignée étroitement avec les besoins d'un type particulier d'avion. Les exigences d'aptitude d'entretien dans ce cas lient au besoin à fournir une organisation de soutien d'entretien qui améliorera l'emploi opérationnel planifié de l'avion, les deux durant peacetime et durant wartime.

### Evaluation Section

Intelligibility: 1 2 3 ④ 5

Accuracy: 0 1 2 3 4 5 ⑥

Comments:

Same remark as before:  
As far as accuracy is concerned, I hesitated between '4' and '6' because I was aware of the English context. If I was a French-speaking reader given the French text, I would not understand at all, not because words or expressions are missing, but solely because of the improper use of words. The problems described in score '3' would fit better, however I cannot say that the content is properly conveyed.

## EVALUATION SHEET 9

Once you have read and understood the preceding instructions, please circle the appropriate score for each scale in the evaluation section.

### English Text

612. While the need to retain capability within the RAAF relates directly to support of operations in the broadest sense, the capability to be established for a particular organizational element must be more closely aligned with the need of a particular aircraft type. Maintenance capability requirements in this case relate to the need to provide a maintenance support organization which will enhance the planned operational employment of the aircraft, both during peacetime and in time of war.

### French Translation

612. Le besoin de retenir l'aptitude dans la Force Royale D'air Australien lie directement au soutien d'opérations dans le sens le plus large. Cependant, l'aptitude à être établie pour un particulier organisationnel élément doit être plus alignée étroitement avec le besoin d'un type particulier d'avion. Les exigences d'aptitude d'entretien dans ce cas lient au besoin à fournir une organisation de soutien d'entretien qui améliorera l'emploi opérationnel planifié de l'avion, les deux, durant des temps de paix et de guerre.

### Evaluation Section

Intelligibility: 1 2 3 ④ 5

Accuracy: 0 1 2 3 ④ 5 6

#### Comments:

While the text is still not clear, and the input not properly conveyed, this version is better than the two previous ones.

## EVALUATION SHEET 10

Once you have read and understood the preceding instructions, please circle the appropriate score for each scale in the evaluation section.

### English Text

304. Flight servicingings comprise those servicingings performed immediately before or after a flight, or between successive flights either to make the aircraft ready for flight or to restore it to a condition to perform further flights. Flight servicingings are OLM tasks and include:

- a. before flight servicingings,
- b. after flight servicingings, and
- c. turnaround servicingings.

### French Translation

304. Les révisions de vol comprennent ces révisions exécutées aussitôt avant ou après un vol, ou entre des vols successifs l'un ou l'autre à faire l'avion prêt pour le vol ou à le restaurer à une condition à exécuter d'avantage vols. Les révisions de vol sont OLM tâches et incluent:

- |    |                        |                                       |
|----|------------------------|---------------------------------------|
| a. | avant vol révisé       | préparation à l'envol                 |
| b. | après le vol révisé et | entretien après-vol                   |
| c. | turnaround révisé      | rotation (préparation au vol suivant) |

### Evaluation Section

Intelligibility: 1 2 3 ④ 5

Accuracy: 0 1 2 3 4 5 ⑥

Comments:

Flight servicingings, for example, should translate into "entretien en piste".

## EVALUATION SHEET 12

Once you have read and understood the preceding instructions, please circle the appropriate score for each scale in the evaluation section.

### English Text

304. Flight servicings comprise those servicings performed immediately before or after a flight, or between successive flights either to make the aircraft ready for flight or to restore it to a condition to perform further flights. Flight servicings are OLM tasks and include:

- a. before flight servicings,
- b. after flight servicings, and
- c. turnaround servicings.

### French Translation

304. Les révisions de vol comprennent ces révisions qui sont exécutées aussitôt avant ou après un vol, ou entre des vols successifs. Les révisions de vol l'un ou l'autre préparent l'avion pour le vol ou le restaurent à une condition à exécuter d'avantage vols. Les révisions de vol sont tâches opérationnelles d'entretien de niveaux et incluent:

- a. avant vol révisé,
- b. après le vol révisé et
- c. la tournure autour des révisions.

### Evaluation Section

Intelligibility: 1 2 3 ④ 5

Accuracy: 0 1 2 3 4 5 ⑥

#### Comments:

Same remark as on sheet 6. I can see an improvement in the choice of words, and 'OLM' needed to be expanded. However, the content is not conveyed at all, and the French expression is improper to give a '4' score. For example, again, I would have translated "Turn-around servicings" into something like "rotation d'appareils". "Tournure autour des révisions" means "turning around reviews". (r)

**Evaluation sheet 15**

313. La procédure de rotation d'appareils (TA) inclut certains éléments des deux procédures de préparation à l'envol (BF) et d'entretien après-vol (AF), mais pas normalement leur totalité. En général, une rotation devrait inclure :

- a. un contrôle préventif d'articles critiques en vue de détecter tout dommage ou toute détérioration qui n'aurait pas été observé(e) par l'équipage durant le vol précédent,
- b. un examen visuel de l'extérieur de l'appareil en vue de constater tout dommage apparent,
- c. réapprovisionnement en carburant si nécessaire.

## EVALUATION SHEET 1

Once you have read and understood the preceding instructions, please circle the appropriate score for each scale in the evaluation section.

### English Text

503. In general, there is no precise method for establishing 'correct' maintenance intervals, because of the inherent variability in equipment performance and lack of complete knowledge about the behaviour of the items concerned. Instead, the process involves setting an interval based on an estimate of the 'correct' value, then progressively refining the estimate by analysis of maintenance and performance data collected for the item.

### French Translation

503. En général, il y a aucune méthode précise pour n'établir 'correct' intervalles d'entretien, à cause de la variabilité inhérente dans (le manque et performance d'équipement de connaissance complète) environ le comportement des articles concernait. Plutôt, le procédé entraîne fixer une intervalle basé sur une estimation du 'correcte' valeur, alors raffinant progressivement l'estimation par l'analyse de données de performance et entretien recueillies pour l'article.

### Evaluation Section

Intelligibility: 1 2 3 4 5

Accuracy: 0 1 2 3 4 5 6

Comments: — machine inconsistent in its ability to recognize the past participle in its role as an adjective following a noun, e.g. "items concerned" ≠ "articles concernait".

"interval based" = "intervalle basé", "data collected" = données ... recueillies

Machine can recognize when past part is followed by prep. phrase, e.g.

"collected for...", "based on ..." etc. Avoid usage of past participle as adj. when it occurs in certain places; e.g. as the final word in the sentence (?)

— I can't figure out why the machine would mix up the order of the words; "in equipment performance and lack of ... concerned." The mixup

"dans le manque ..." really plays havoc with the sentence's meaning.

— Sentence beginning "Instead ..." is too long; machine fails to recognize connection between "involves setting" and "then" involves "progressively refining." Perhaps the English text needs to repeat words that are only 'understood' by native speakers, e.g. the second "involves".

- machine modifies 'correct' to 'correcte' but fails to change "du" to "de la"
- machine is inconsistent in making past participles agree with the nouns they modify e.g. "un intervalle basé" - no agreement  
"données ... recueillies" - agreement

## EVALUATION SHEET 2

Once you have read and understood the preceding instructions, please circle the appropriate score for each scale in the evaluation section.

### English Text

503. In general, there is no precise method for establishing 'correct' maintenance intervals, because of the inherent variability in equipment performance and lack of complete knowledge about the behaviour of the items concerned. Instead, the process involves setting an interval based on an estimate of the 'correct' value, then progressively refining the estimate by analysis of maintenance and performance data collected for the item.

### French Translation

503. En général, il y a aucune méthode précise pour n'établir 'correct' intervalles d'entretien à cause de la variabilité inhérente dans la performance d'équipement et le manque de connaissance complète environ le comportement des articles concernait. Plutôt, le procédé commence en fixant une intervalle basé sur une estimation du 'correcte' valeur. (Cela estimait l'intervalle) peut alors être raffiné progressivement en analysant des données de performance et entretien recueillies pour l'article.

### Evaluation Section

Intelligibility: 1 2 3 4 5

Accuracy: 0 1 2 3 4 5 6

Comments: Machine's attempt to break up the too-long English sentence runs into difficulty e.g. Cela "estimait l'intervalle." Presumably it wanted to translate: "That estimate of the interval," and presumably it ran into problems differentiating between "that" as demonstrative pronoun and "that" as an adjective; and between "estimate" as a noun and "estimate" as a verb. On second thought, if the machine was trying to translate "That estimated interval" then "cela estimait l'intervalle" would be the logical outcome of its failure to recognize "estimated" as ~~an~~ participle/adjective instead of a past tense verb.

Computer needs to be programmed somehow to handle the very tricky problem of the translation of prepositions. For example, the



preposition "about" can mean "concerning" or "approximately,"

Either the English text must be very explicit, seeking to eliminate all ambiguity, or the computer must have some examples with which to compare the text at hand. Thus, when it sees "about the behavior," it will check that phrase against its own store of possibilities, e.g.

"about fifteen" = "environs<sup>(15)</sup> quinze", "about my friend" = "(de) mon ami." It will, therefore, eliminate "environs" as a possibility.

### EVALUATION SHEET 3

Once you have read and understood the preceding instructions, please circle the appropriate score for each scale in the evaluation section.

#### English Text

503. In general, there is no precise method for establishing 'correct' maintenance intervals, because of the inherent variability in equipment performance and lack of complete knowledge about the behaviour of the items concerned. Instead, the process involves setting an interval based on an estimate of the 'correct' value, then progressively refining the estimate by analysis of maintenance and performance data collected for the item.

#### French Translation

503. Typiquement, il y a aucune méthode précise pour n'établir 'correct' intervalles d'entretien. Cette difficulté survient à cause de la variabilité inhérente dans la performance d'équipement et le manque de connaissance complète environ le comportement des articles concernait. Plutôt, le procédé entraîne fixer une intervalle qui est basée sur une estimation du 'correcte' valeur. Cela estimait l'intervalle est alors raffiné progressivement par une analyse des données de performance et entretien recueillies pour l'article.

#### Evaluation Section

Intelligibility: 1 2 3 4 5

Accuracy: 0 1 2 3 4 5 6

Comments: Except for "concernait" and "Cela estimait l'intervalle," I find this translation the easiest to understand. Though it alters the English <sup>sentence structure</sup> more radically than the other 2 translations, it manages to convey more accurately the essence of the English text.

#### EVALUATION SHEET 4

Once you have read and understood the preceding instructions, please circle the appropriate score for each scale in the evaluation section.

#### English Text

521. Review of Routine Servicing Intervals. The routine servicing intervals for an aircraft are always reviewed as a set, in conjunction with a review of the tasks performed. Thus, the review task is essentially one of determining the continuing validity and effectiveness of the tasks performed on each item included in the set of routine servicings, and of the intervals at which they are performed. In addition the requirements must be considered in conjunction with flight servicing and DLM requirements, to ensure that each task on each item is performed at a single appropriate interval.

#### French Translation

521. La revue de Routine Révisant Des intervalles. La routine révisant des intervalles pour un avion sont toujours revus comme un ensemble, dans la conjonction avec une revue des tâches exécutées. Ainsi, la tâche de revue est essentiellement une de déterminer l'efficacité et validité continue des tâches exécutée sur chaque article incluse dans l'ensemble de révisions de routine, et des intervalles à lesquelles ils sont exécutés. Au surplus les exigences doivent être considérées dans la conjonction avec le vol révisant et DLM exigences, à assurer que chaque tâche sur chaque article est exécutée à une intervalle appropriée seule.

#### Evaluation Section

Intelligibility: 1 2 3 4 5

Accuracy: 0 1 2 3 4 5 6

Comments: I'm not surprised the machine had trouble with "Routine Servicing Intervals." How is it supposed to know that routine is not a noun in this instance, but rather, an adjective? Perhaps if "Servicing Intervals" were hyphenated, then the computer would recognize that "routine" would <sup>as is</sup> have to be an adjective. The other solution would be to leave such expressions as "routine servicing intervals", that is, adj/noun + "ing" word + noun. <sup>Then</sup> To avoid confusion all English texts would have to be free of all similar-looking but differently-functioning constructions such as "Coat requiring attention" or "a quote needing verification," e.g. all <sup>verb/noun + "ing" word + noun constructions</sup> Eng text should, where possible, use the most explicit form when using prepositions. ~~So~~ A prepositional such as "to" in "to ensure" should be changed to "in order to" so that the French translation becomes "pour" in -

Steas of "à", or "afin de" instead of "à".

## EVALUATION SHEET 5

Once you have read and understood the preceding instructions, please circle the appropriate score for each scale in the evaluation section.

### English Text

521. Review of Routine Servicing Intervals. The routine servicing intervals for an aircraft are always reviewed as a set, in conjunction with a review of the tasks performed. Thus, the review task is essentially one of determining the continuing validity and effectiveness of the tasks performed on each item included in the set of routine servicings, and of the intervals at which they are performed. In addition the requirements must be considered in conjunction with flight servicing and DLM requirements, to ensure that each task on each item is performed at a single appropriate interval.

### French Translation

521. La revue de Routine Révisant Des intervalles. La routine révisant des intervalles pour un avion sont toujours revus comme un ensemble ~~ensemble~~ avec une revue des tâches exécutés. Ainsi, la tâche de revue est essentiellement une de déterminer l'efficacité et validité continue des deux les tâches exécutée sur chaque article incluse dans l'ensemble de révisions de routine, et les intervalles à lesquelles les tâches sont exécutées. Au surplus, les exigences doivent être considérées ensemble avec le vol révisant et DLM exigences à assurer que chaque tâche sur chaque article est exécutée à une intervalle appropriée seule.

### Evaluation Section

Intelligibility: 1 2 3 4 5

Accuracy: 0 1 2 3 4 5 6

Comments: "des deux les tâches" — presumably the machine was trying to translate "of both the tasks .... and." It <sup>also</sup> knows that "of both tasks" is "des deux tâches;" it <sup>also</sup> knows that the plural definite article is "les." Hence, "des deux les tâches." It doesn't know that it should be translating a "both — and —" construction because there is too long a gap between the both and elements, e.g. between "tasks" and "and."

## EVALUATION SHEET 6

Once you have read and understood the preceding instructions, please circle the appropriate score for each scale in the evaluation section.

### English Text

521. Review of Routine Servicing Intervals. The routine servicing intervals for an aircraft are always reviewed as a set, in conjunction with a review of the tasks performed. Thus, the review task is essentially one of determining the continuing validity and effectiveness of the tasks performed on each item included in the set of routine servicings, and of the intervals at which they are performed. In addition the requirements must be considered in conjunction with flight servicing and DLM requirements, to ensure that each task on each item is performed at a single appropriate interval.

### French Translation

521. La revue de Routine Révisant Des intervalles. La routine révisant des intervalles pour un avion sont toujours revus comme un ensemble, dans la conjonction avec une revue des tâches qui sont exécutés. La tâche de revue est essentiellement une de déterminer l'efficacité et validité continue des tâches exécutées sur chaque article incluse dans l'ensemble de révisions de routine. La revue examine alors les intervalles à lesquelles tâches valables pour chaque article sont exécutés. Au surplus, les exigences doivent être considérées dans la conjonction avec la révision de vol et Niveau De dépôt exigences D'entretien. Cette considération assurera que toutes tâches pour chaque article sont exécutées à une intervalle appropriée seule.

### Evaluation Section

Intelligibility: 1 2 (3) 4 5

Accuracy: 0 1 2 3 (4) 5 6

#### Comments:

- "Depot Level Maintenance Requirements" = should be reworded to say (noun/adj + noun/adj + noun/adj + noun) "the requirements of DLH (Depot Level Main.)"  
- other words, reduce the number of noun/adj. words that form any one phrase without an intervening word.  
All such constructions in English should be similarly worded.
- See back of page

<u>TEXT</u>	Vs	MORE GRAMMATICALLY CORRECT (?) VERSION OR BETTER (?) VERSION
dans la conjonction	Vs	en conjonction
comme un ensemble	Vs	en groupe
sont exécutés	Vs	ont été exécutés
l'efficacité et validité	Vs	l'efficacité et <u>la</u> validité
continue	Vs	continues
exécutée	Vs	exécutées
à lesquelles tâches	Vs	<u>auxquelles</u> <u>les</u> tâches
au surplus	Vs	en plus
DLM (Depot Level Maintenance) ?	Vs	the requirements of Depot Level Maintenance
Niveau de dépôt exigences d'entretien	Vs	<u>les</u> exigences d'entretien au niveau dépôt (?) (de dépôt)
toutes tâches	Vs	toutes <u>les</u> tâches
sont exécutées	Vs	<u>seront</u> exécutées (to agree with future) "assurera"
une intervalle appropriée	Vs	une <u>seule</u> intervalle appropriée

## EVALUATION SHEET 7

Once you have read and understood the preceding instructions, please circle the appropriate score for each scale in the evaluation section.

### English Text

612. While the need to retain capability within the RAAF relates directly to support of operations in the broadest sense, the capability to be established for a particular organizational element must be more closely aligned with the need of a particular aircraft type. Maintenance capability requirements in this case relate to the need to provide a maintenance support organization which will enhance the planned operational employment of the aircraft, both during peacetime and in time of war.

### French Translation

612. Tandis que le besoin de retenir l'aptitude dans le RAAF lie soutenir directement d'opérations dans le sens le plus large, l'aptitude à être établie pour un particulier organizationnel élément doit être plus alignée étroitement avec le besoin d'un type particulier d'avion. Les exigences d'aptitude d'entretien dans ce cas lient au besoin à fournir une organisation de soutien d'entretien qui améliorera l'emploi opérationnel planifié de l'avion, les deux durant peacetime et à temps de guerre.

### Evaluation Section

Intelligibility: 1 2 3 4 5

Accuracy: 0 1 2 3 4 5 6

#### Comments:

- Care with placement of adverbs, especially if between a verb & infinitive, e.g. "directly relates to support" versus "relates directly to support" will prevent the computer from seeing "to support", in this instance, as an infinitive. Another way to prevent the computer from making the error, in this instance, would be to change the English text to "relates directly to the support."
- Placement of "both" in the English text from "both during" to "during both" would prevent the grammatically correct, but contextually incorrect, "les deux."



## EVALUATION SHEET 8

Once you have read and understood the preceding instructions, please circle the appropriate score for each scale in the evaluation section.

### English Text

612. While the need to retain capability within the RAAF relates directly to support of operations in the broadest sense, the capability to be established for a particular organizational element must be more closely aligned with the need of a particular aircraft type. Maintenance capability requirements in this case relate to the need to provide a maintenance support organization which will enhance the planned operational employment of the aircraft, both during peacetime and in time of war.

### French Translation

612. Le besoin de retenir l'aptitude d'entretien dans le RAAF lie directement au soutien d'opérations d'avion dans le sens le plus large. Cependant, l'aptitude d'entretien à être établie pour un particulier organizationnel élément doit être alignée étroitement avec les besoins d'un type particulier d'avion. Les exigences d'aptitude d'entretien dans ce cas lient au besoin à fournir une organisation de soutien d'entretien qui améliorera l'emploi opérationnel planifié de l'avion, les deux durant peacetime et durant wartime.

### Evaluation Section

Intelligibility: 1 2 3 4 5

Accuracy: 0 1 2 3 4 5 6

Comments:

- Since "Capability" refers to "intellectual" as well as other ability, the Computer must have some way of eliminating "aptitude" as a translation in this paragraph.
- "Maintenance capability requirements" doesn't confuse the Computer as did "Dept Level Maintenance Requirements" on sheets 4, 5 & 6

## EVALUATION SHEET 9

Once you have read and understood the preceding instructions, please circle the appropriate score for each scale in the evaluation section.

### English Text

612. While the need to retain capability within the RAAF relates directly to support of operations in the broadest sense, the capability to be established for a particular organizational element must be more closely aligned with the need of a particular aircraft type. Maintenance capability requirements in this case relate to the need to provide a maintenance support organization which will enhance the planned operational employment of the aircraft, both during peacetime and in time of war.

### French Translation

612. Le besoin de retenir l'aptitude dans la Force Royale D'air Australien lie directement au soutien d'opérations dans le sens le plus large. Cependant, l'aptitude à être établie pour un particulier organizational élément doit être plus alignée étroitement avec le besoin d'un type particulier d'avion. Les exigences d'aptitude d'entretien dans ce cas lient au besoin à fournir une organisation de soutien d'entretien qui améliorera l'emploi opérationnel planifié de l'avion, les deux durant des temps de paix et de guerre.

### Evaluation Section

Intelligibility: 1 (2) 3 4 5

Accuracy: 0 1 2 (3) 4 5 6

Comments:

Computer needs to be able to recognize that "more" is an adverb as well as an adjective. It needs to know that adjective + adverb + ~~past participle~~ is not a possibility in English and that "plus" in French stays with the adverb it modifies.

## EVALUATION SHEET 10

Once you have read and understood the preceding instructions, please circle the appropriate score for each scale in the evaluation section.

### English Text

304. Flight servicings comprise those servicings performed immediately before or after a flight, or between successive flights either to make the aircraft ready for flight or to restore it to a condition to perform further flights. Flight servicings are OLM tasks and include:

- a. before flight servicings,
- b. after flight servicings, and
- c. turnaround servicings.

### French Translation

304. Les révisions de vol comprennent ces révisions exécutaient aussitôt avant ou après un vol, ou entre des vols successifs l'un ou l'autre à faire l'avion prêt pour le vol ou à le restaurer à une condition à exécuter davantage vols. Les révisions de vol sont OLM tâches et incluent:

- a. avant vol révisé,
- b. après le vol révisé, et
- c. turnaround révisé.

### Evaluation Section

Intelligibility: 1 2 3 4 5

Accuracy: 0 1 2 3 4 5 6

#### Comments:

- "those servicings performed" might be phrased "the servicings which are performed." Thus the error "ces révisions exécutaient," would be eliminated.
- Computer might be programmed to scan a sentence for expressions such as "to make ready" in order to avoid a word-for-word translation.
- Computer can't recognize "before" as an element of "before flight"; perhaps a hyphen should join the two, e.g. "before-flight".
- Computer doesn't recognize "servicings" as a <sup>plural</sup> noun form as opposed to a third person sing. verb form.

## EVALUATION SHEET 11

Once you have read and understood the preceding instructions, please circle the appropriate score for each scale in the evaluation section.

### English Text

304. Flight servicings comprise those servicings performed immediately before or after a flight, or between successive flights either to make the aircraft ready for flight or to restore it to a condition to perform further flights. Flight servicings are OLM tasks and include:

- a. before flight servicings,
- b. after flight servicings, and
- c. turnaround servicings.

### French Translation

304. Les révisions de vol sont ces révisions exécutaient aussitôt avant ou après un vol, ou entre des vols successifs. Ils sont exécutés à l'un ou l'autre faits l'avion prêt pour le vol ou restaurés à un flyable condition. Les révisions de vol sont OLM tâches. Ils incluent avant révisions de vol, après des révisions de vol, et turnaround révisions.

### Evaluation Section

Intelligibility: 1 2 3 (4) 5

Accuracy: 0 1 2 3 (4) 5 6

Comments:

The Computer recognizes "before" and "after" as prepositions and it groups "flight servicings" as a single element. Again, some way needs to be found to let the Computer know that the words to be considered as a single element are: "before flight."

### EVALUATION SHEET 13

Once you have read and understood the preceding instructions, please circle the appropriate score for each scale in the evaluation section.

#### English Text

313. The TA servicing comprises elements of both the BF and AF servicings, but does not normally include all requirements of each servicing. In general a TA servicing should include:

- a. checks of safety critical items for damage or deterioration which would not have been evident to the crew during the preceding flight,
- b. a visual examination of the exterior of the aircraft for obvious damage, and
- c. replenishment as required.

#### French Translation

313. Le TA révisant comprend des éléments des deux le BF et AF révisions, mais n'inclut pas normalement toutes exigences de chacun révisant. En général un TA révision devrait inclure:

- a. les chèques de sécurité articles critiques pour le dommages ou détérioration que n'aurait pas été évidente à l'équipage durant le vol précédent,
- b. un examen visuel de l'extérieur de l'avion pour le dommages évidents, et
- c. replenishment comme requis.

#### Evaluation Section

Intelligibility: 1 2 3 4 5

Accuracy: 0 1 2 3 4 5 6

#### Comments:

- Computer needs to be programmed to recognize in a series of words such as noun/adj. + adj. + noun, which words belong together, e.g. "safety critical"
- computer needs to know that if the Eng is sing. but the French is plural, the article in French must reflect the change, e.g. singular "damage" becomes plural "dommages." Therefore, the article

"le" must become "les"

- Computer needs to recognize that "each" may be either a pronoun or adjective and to differentiate between the two.

#### EVALUATION SHEET 14

Once you have read and understood the preceding instructions, please circle the appropriate score for each scale in the evaluation section.

#### English Text

313. The TA servicing comprises elements of both the BF and AF servicings, but does not normally include all requirements of each servicing. In general a TA servicing should include:

- a. checks of safety critical items for damage or deterioration which would not have been evident to the crew during the preceding flight,
- b. a visual examination of the exterior of the aircraft for obvious damage, and
- c. replenishment as required.

#### French Translation

313. Le TA révisant inclut des éléments des deux le BF et AF révisions, mais n'inclut pas normalement toutes les exigences de chacun révisant. En général, un TA révision devrait inclure:

les chèques de sécurité articles critiques pour le dommages ou détérioration que n'aurait pas été évidente à l'équipage durant le vol précédent,

un examen visuel de l'extérieur de l'avion pour le dommages évident, et

le replenishment de fluides et gazer si nécessaires.

#### Evaluation Section

Intelligibility: 1 2 3 (4) 5

Accuracy: 0 1 2 3 (4) 5 6

#### Comments:

Computer needs to recognize the function of "which" in a sentence, e.g. In which would not have been evident " , "which" is the subject of the clause and therefore should be translated by "qui" rather than "que".

## EVALUATION SHEET 15

Once you have read and understood the preceding instructions, please circle the appropriate score for each scale in the evaluation section.

### English Text

313. The TA servicing comprises elements of both the BF and AF servicings, but does not normally include all requirements of each servicing. In general a TA servicing should include:

- a. checks of safety critical items for damage or deterioration which would not have been evident to the crew during the preceding flight,
- b. a visual examination of the exterior of the aircraft for obvious damage, and
- c. replenishment as required.

### French Translation

313. La Tournure Autour réviser inclut des éléments des deux L'avant Vol et Après des révisions De vol, mais il n'inclut pas normalement toutes les exigences de chacun révisant. Typiquement une Tournure Autour la révision devrait inclure:

- a. les chèques d'articles critiques de sécurité pour le ~~dommages~~ ou ~~détérioration~~ que n'aurait pas été évidente à l'équipage durant le vol précédent,
- b. un examen visuel de l'extérieur de l'avion pour le ~~dommages~~ évident, et
- c. replenishment comme requis.

### Evaluation Section

Intelligibility: 1 2 3 (4) 5

Accuracy: 0 1 2 3 (4) 5 6

#### Comments:

Computer correctly translates "Before flight" as "avant vol" but doesn't realize that the word "servicings" is understood. Perhaps the word "servicings" should be used after "BF" as well as after "AF". Computer correctly <sup>translates</sup> ~~identifies~~ "after flight servicings" as "après des révisions de vol." However, in this context, the translation is incorrect. Again, a signal needs to be sent to the computer that "before flight" acts as an adjectival phrase.



18 August 1992

General comments on NCR machine translation evaluation:

1. The criteria listed for the scoring gradations seemed somewhat confusing. Specifically, for "accuracy," there seems to be too much stress on "missing" elements in nos. 4-6. In fact, none of the translations had anything "missing" in this sense. However, one wants to have a rating category that applies to problems that are more severe than those listed for no. 3. So I used no. 4 for such severe problems, despite the fact that nothing was "missing." Indeed, if the fact of "missing" elements did not seem to gain in importance as the scale progresses from 4 to 6, I might have used 5 or 6 rather than 3 or 4 in some of my ratings.

2. In general, with machine translation the results are much worse when the desired "transfers" (translations) for field-specific terms and phrases have not been adequately entered into the lexicon. When the machine uses "general" meanings for technical or field-specific terms an evaluator will be distracted by the resulting "howlers" of mistranslation and tend to overlook the grammatical problems that are more indicative of genuine faults of the translation system. In this evaluation, it was striking that almost none of the variants seemed designed to show how the system worked when all the relevant terminology had been entered. In this sense, I think the test was unfair to the machine. Most of the variants seemed more to have been slightly doctored by human editing than to reflect differences in machine performance based on different parameter settings or rules put into effect for the specific translation run.

## EVALUATION SHEET 1

Once you have read and understood the preceding instructions, please circle the appropriate score for each scale in the evaluation section.

### English Text

503. In general, there is no precise method for establishing 'correct' maintenance intervals, because of the inherent variability in equipment performance and lack of complete knowledge about the behaviour of the items concerned. Instead, the process involves setting an interval based on an estimate of the 'correct' value, then progressively refining the estimate by analysis of maintenance and performance data collected for the item.

### French Translation

503. En général, il y a aucune méthode précise pour n'établir 'correct' intervalles d'entretien, à cause de la variabilité inhérente dans le manque et performance d'équipement de connaissance complète environ le comportement des articles concernait. Plutôt, le procédé entraîne fixer une intervalle basé sur une estimation du 'correcte' valeur, alors raffinant progressivement l'estimation par l'analyse de données de performance et entretien recueillies pour l'article.

### Evaluation Section

Intelligibility: 1 2 3 (4) 5

Accuracy: 0 1 2 3 (4) 5 6

#### Comments:

The machine was unable to grasp the parallelism between 'variability' and 'lack' — the two components of the noun phrase would have been more clearly separated by "as well as" rather than "and".

#### EVALUATION SHEET 4

Once you have read and understood the preceding instructions, please circle the appropriate score for each scale in the evaluation section.

#### English Text

521. Review of Routine Servicing Intervals. The routine servicing intervals for an aircraft are always reviewed as a set, in conjunction with a review of the tasks performed. Thus, the review task is essentially one of determining the continuing validity and effectiveness of the tasks performed on each item included in the set of routine servicings, and of the intervals at which they are performed. In addition the requirements must be considered in conjunction with flight servicing and DLM requirements, to ensure that each task on each item is performed at a single appropriate interval.

#### French Translation

521. La revue de Routine Révisant Des intervalles. La routine révisant des intervalles pour un avion sont toujours revus comme un ensemble, dans la conjonction avec une revue des tâches exécutés. Ainsi, la tâche de revue est essentiellement une de déterminer l'efficacité et validité continue des tâches exécutée sur chaque article incluse dans l'ensemble de révisions de routine, et des intervalles à lesquelles ils sont exécutés. Au surplus les exigences doivent être considérées dans la conjonction avec le vol révisant et DLM exigences, à assurer que chaque tâche sur chaque article est exécutée à une intervalle appropriée seule.

#### Evaluation Section

Intelligibility: 1 2 3 4 (5)

Accuracy: 0 1 2 3 (4) 5 6

#### Comments:

"ing" always cause severe problems for machine  
suggestion it would be better to use the noun "service".

## EVALUATION SHEET 1

Once you have read and understood the preceding instructions, please circle the appropriate score for each scale in the evaluation section.

### English Text

503. In general, there is no precise method for establishing 'correct' maintenance intervals, because of the inherent variability in equipment performance and lack of complete knowledge about the behaviour of the items concerned. Instead, the process involves setting an interval based on an estimate of the 'correct' value, then progressively refining the estimate by analysis of maintenance and performance data collected for the item.

### French Translation

503. En général, il y a aucune méthode précise pour n'établir 'correct' intervalles d'entretien, à cause de la variabilité inhérente dans le manque et performance d'équipement de connaissance complète environ le comportement des articles concernait. Plutôt, le procédé entraîne fixer une intervalle basé sur une estimation du 'correcte' valeur, alors raffinant progressivement l'estimation par l'analyse de données de performance et entretien recueillies pour l'article.

### Evaluation Section

Intelligibility: 1 2 3 4 5 .

Accuracy: 0 1 2 3 4 5 6

Comments:

SHORTEN SENTENCES IN ENGLISH TEXT  
may have helped; also if grammatically simpler.

### EVALUATION SHEET 3

Once you have read and understood the preceding instructions, please circle the appropriate score for each scale in the evaluation section.

#### English Text

503. In general, there is no precise method for establishing 'correct' maintenance intervals, because of the inherent variability in equipment performance and lack of complete knowledge about the behaviour of the items concerned. Instead, the process involves setting an interval based on an estimate of the 'correct' value, then progressively refining the estimate by analysis of maintenance and performance data collected for the item.

#### French Translation

503. Typiquement, il y a aucune méthode précise pour n'établir 'correct' intervalles d'entretien. Cette difficulté survient à cause de la variabilité inhérente dans la performance d'équipement et le manque de connaissance complète environ le comportement des articles concernés. Plutôt, le procédé entraîne fixer une intervalle qui est basée sur une estimation du 'correcte' valeur. Cela estimait l'intervalle est alors raffiné progressivement par une analyse des données de performance et entretien recueillies pour l'article.

#### Evaluation Section

Intelligibility: 1 2 3 (4) 5

Accuracy: 0 1 2 3 (4) 5 6

Comments: See evaluation sheet 1.

The machine split the first English sentence into 2 French sentences, and inserted a period between these two.

#### EVALUATION SHEET 4

Once you have read and understood the preceding instructions, please circle the appropriate score for each scale in the evaluation section.

#### English Text

521. Review of Routine Servicing Intervals. The routine servicing intervals for an aircraft are always reviewed as a set, in conjunction with a review of the tasks performed. Thus, the review task is essentially one of determining the continuing validity and effectiveness of the tasks performed on each item included in the set of routine servicings, and of the intervals at which they are performed. In addition the requirements must be considered in conjunction with flight servicing and DLM requirements, to ensure that each task on each item is performed at a single appropriate interval.

#### French Translation

521. La revue de Routine Révisant Des intervalles. La routine révisant des intervalles pour un avion sont toujours revus comme un ensemble, dans la conjonction avec une revue des tâches exécutés. Ainsi, la tâche de revue est essentiellement une de déterminer l'efficacité et validité continue des tâches exécutée sur chaque article incluse dans l'ensemble de révisions de routine, et des intervalles à lesquelles ils sont exécutés. Au surplus les exigences doivent être considérées dans la conjonction avec le vol révisant et DLM exigences, à assurer que chaque tâche sur chaque article est exécutée à une intervalle appropriée seule.

#### Evaluation Section

Intelligibility: 1 2 3 4 5  
Accuracy: 0 1 2 3 4 5 6

Comments: English sentences too long & complex.

## EVALUATION SHEET 8

Once you have read and understood the preceding instructions, please circle the appropriate score for each scale in the evaluation section.

### English Text

612. While the need to retain capability within the RAAP relates directly to support of operations in the broadest sense, the capability to be established for a particular organizational element must be more closely aligned with the need of a particular aircraft type. Maintenance capability requirements in this case relate to the need to provide a maintenance support organization which will enhance the planned operational employment of the aircraft, both during peacetime and in time of war.

### French Translation

612. Le besoin de retenir l'aptitude d'entretien dans le RAAP lie directement au soutien d'opérations d'avion dans le sens le plus large. Cependant, l'aptitude d'entretien à être établie pour un particulier organizational élément doit être alignée étroitement avec les besoins d'un type particulier d'avion. Les exigences d'aptitude d'entretien dans ce cas lient au besoin à fournir une organisation de soutien d'entretien qui améliorera l'emploi opérationnel planifié de l'avion, les deux durant peacetime et durant wartime.

### Evaluation Section

Intelligibility: 1 2 3 (4) 5

Accuracy: 0 1 2 3 4 (5) 6

Comments: See Evaluation Sheet 3.

The first English sentence has been split into two French sentences separated by a period

## EVALUATION SHEET 9

Once you have read and understood the preceding instructions, please circle the appropriate score for each scale in the evaluation section.

### English Text

612. While the need to retain capability within the RAAF relates directly to support of operations in the broadest sense, the capability to be established for a particular organizational element must be more closely aligned with the need of a particular aircraft type. Maintenance capability requirements in this case relate to the need to provide a maintenance support organization which will enhance the planned operational employment of the aircraft, both during peacetime and in time of war.

### French Translation

612. Le besoin de retenir l'aptitude dans la Force Royale D'air Australien lie directement au soutien d'opérations dans le sens le plus large. Cependant, l'aptitude à être établie pour un particulier organizational élément doit être plus alignée étroitement avec le besoin d'un type particulier d'avion. Les exigences d'aptitude d'entretien dans ce cas lient au besoin à fournir une organisation de soutien d'entretien qui améliorera l'emploi opérationnel planifié de l'avion, les deux durant des temps de paix et de guerre.

### Evaluation Section

Intelligibility: 1 2 3 (4) 5

Accuracy: 0 1 2 3 4 (5) 6

Comments: See previous evaluation sheets.

I doubt that the underlined (by me) phrase was machine-translated, unless it was entered as a single idiom in the dictionary.



## EVALUATION SHEET 12

Once you have read and understood the preceding instructions, please circle the appropriate score for each scale in the evaluation section.

### English Text

304. Flight servicings comprise those servicings performed immediately before or after a flight, or between successive flights either to make the aircraft ready for flight or to restore it to a condition to perform further flights. Flight servicings are OLM tasks and include:

- a. before flight servicings,
- b. after flight servicings, and
- c. turnaround servicings.

### French Translation

304. Les révisions de vol comprennent ces révisions qui sont exécutées aussitôt avant ou après un vol, ou entre des vols successifs. Les révisions de vol l'un ou l'autre préparent l'avion pour le vol ou le restaurent à une condition à exécuter davantage vols. Les révisions de vol sont tâches opérationnelles D'entretien De niveau et incluent:

- a. avant vol révisé,
- b. après le vol révisé, et
- c. la tournure autour des révisions.

### Evaluation Section

Intelligibility: 1 2 3 4 5

Accuracy: 0 1 2 3 4 5 6

#### Comments:

The first English sentence has been split into two French sentences separated by a period

Appendix F: Graphs of Frequency Distributions of Scores

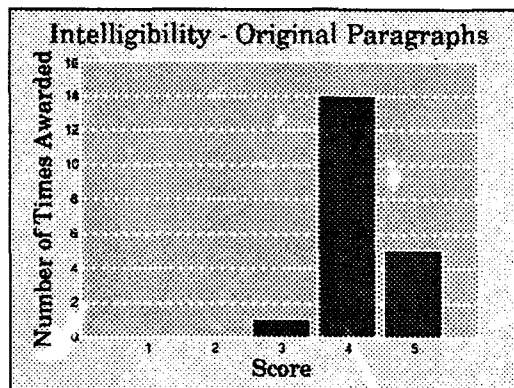


Fig. 2. Frequency Distribution for Intelligibility - Original Paragraphs

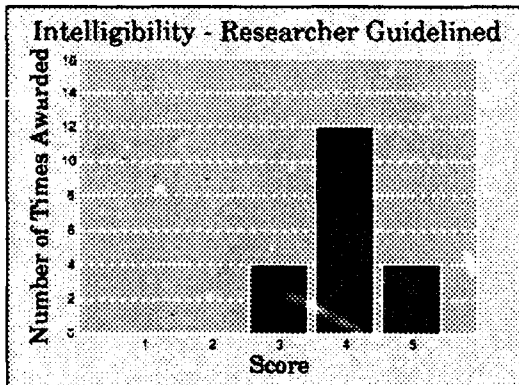


Fig. 3. Frequency Distribution for Intelligibility - Researcher Guidelined Paragraphs

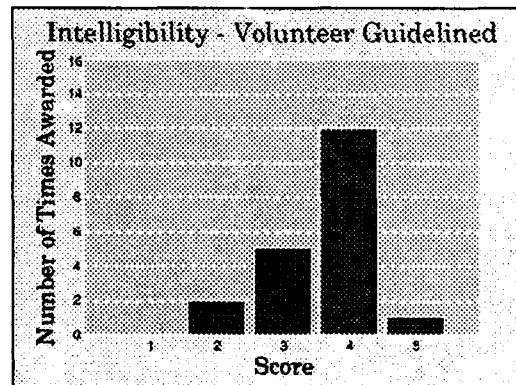


Fig. 4. Frequency Distribution for Intelligibility - Volunteer Guidelined Paragraphs

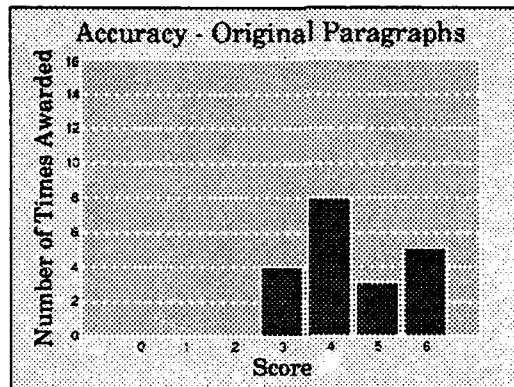


Fig. 5. Frequency Distribution for Accuracy - Original Paragraphs

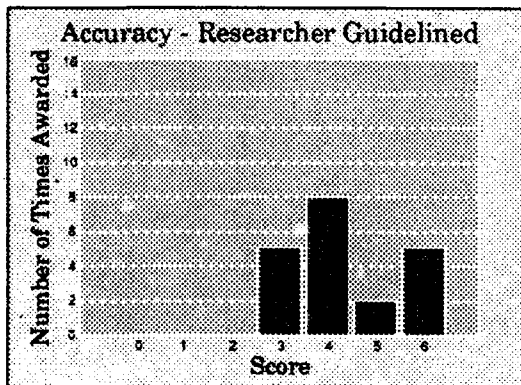


Fig. 6. Frequency Distribution for Accuracy - Researcher Guidelined Paragraphs

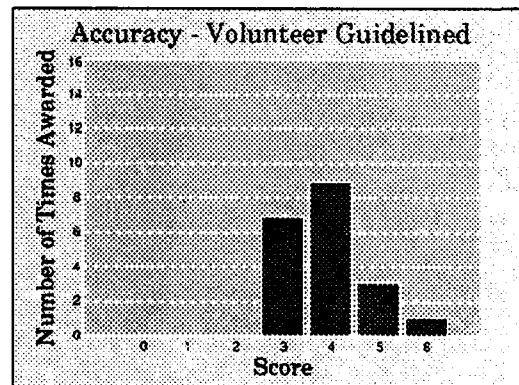


Fig. 7. Frequency Distribution for Accuracy - Volunteer Guidelined Paragraphs

Appendix G: Statistix Sign Test Reports

Report for Intelligibility of Researcher Guidelined Paragraphs vs Originals

SIGN TEST FOR INT\_ORIG - INT\_RES

NUMBER OF NEGATIVE DIFFERENCES	2
NUMBER OF POSITIVE DIFFERENCES	6
NUMBER OF ZERO DIFFERENCES (IGNORED)	12

PROBABILITY OF A RESULT AS  
OR MORE EXTREME THAN OBSERVED 0.1445

A VALUE IS COUNTED AS A ZERO IF ITS  
ABSOLUTE VALUE IS LESS THAN 0.00001

CASES INCLUDED 20 MISSING CASES 0

Report for Intelligibility of Volunteer Guidelined Paragraphs vs Originals

SIGN TEST FOR INT\_ORIG - INT\_VOL

NUMBER OF NEGATIVE DIFFERENCES	0
NUMBER OF POSITIVE DIFFERENCES	11
NUMBER OF ZERO DIFFERENCES (IGNORED)	9

PROBABILITY OF A RESULT AS  
OR MORE EXTREME THAN OBSERVED 0.0005

A VALUE IS COUNTED AS A ZERO IF ITS  
ABSOLUTE VALUE IS LESS THAN 0.00001

CASES INCLUDED 20 MISSING CASES 0

Report for Accuracy of Researcher Guidelined Paragraphs vs Originals

SIGN TEST FOR ACC\_ORIG - ACC\_RES

NUMBER OF NEGATIVE DIFFERENCES	3
NUMBER OF POSITIVE DIFFERENCES	4
NUMBER OF ZERO DIFFERENCES (IGNORED)	13

PROBABILITY OF A RESULT AS  
OR MORE EXTREME THAN OBSERVED 0.5000

A VALUE IS COUNTED AS A ZERO IF ITS  
ABSOLUTE VALUE IS LESS THAN 0.00001

CASES INCLUDED 20 MISSING CASES 0

Report for Accuracy of Volunteer Guidelined Paragraphs vs Originals

SIGN TEST FOR ACC\_ORIG - ACC\_VOL

NUMBER OF NEGATIVE DIFFERENCES	0
NUMBER OF POSITIVE DIFFERENCES	8
NUMBER OF ZERO DIFFERENCES (IGNORED)	12

PROBABILITY OF A RESULT AS  
OR MORE EXTREME THAN OBSERVED 0.0039

A VALUE IS COUNTED AS A ZERO IF ITS  
ABSOLUTE VALUE IS LESS THAN 0.00001

CASES INCLUDED 20 MISSING CASES 0

### Bibliography

1. "AI Techniques are Being Applied to Language Translation," Research and Development, 30: 67-68 (April 1988).
2. Akiyama, Ryoza. "Our Experience in Using SYSTRAN," Machine Translation Summit. 150-151. Tokyo: Ohmsha, Ltd., 1989.
3. Aldridge, 1 Lt Lee F. French-English Machine Translation - General Preliminaries and Lexical Routine Development. MS Thesis AFIT/GE/EE/72-1. School of Engineering, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, December 1971 (AD-741259).
4. Alonso, Juan A. "Machine Translation Technology: On the Way to Market Introduction," Siemens Review, 54: 32-36 (November 1987).
5. Amano, Shin-ya et al. "AS-TRANSAC: The Toshiba Translation System," Machine Translation Summit. 107-112. Tokyo: Ohmsha, Ltd., 1989.
6. ----- "The Toshiba Machine Translation System," Future Generations Computer Systems, 2: 121-123 (1986).
7. Arnold, Doug. "Eurotra: A European Perspective on MT," Proceedings of the IEEE, 74: 979-992 (July 1986).
8. ----- and R.L. Johnson. "Robust Processing in Machine Translation," Proceedings of COLING 84. 472-475. Stanford CA, 1984.
9. Bateman, Russell. "Introduction to Interactive Translation," Tools for the Trade, Translating and the Computer 5. 193-197. London: Aslib, 1985.
10. Benton, Perter M. "The Multilingual Edge," Byte, 16: 124-132 (March 1991).
11. Boschen, A.C. "Esperanto, to Simplify and Clarify International Communication," 1989 IEEE International Professional Communication Conference. 71-76. IEEE Press, 1989.
12. Bostad, Dale A. "Aspects of Machine Translation in the United States Air Force," AGARD Lecture Series No. 171, Benefits of Computer Assisted Translation to Information Managers and End-Users. 6-1-6-5. Neuilly sur Seine France: AGARD, 1990 (AD-A225 965).
13. Buican, I.G. and M Amador. "International English," 1990 IEEE International Professional Communication Conference. 217. IEEE Press, 1990.

14. Calistro, Ralph F. "People and Machines: An Experience in Machine-Aided Translation," Proceedings of the International Technical Communication Conference. RT58-RT61. 1991.
15. Cambridge Encyclopedia, The. Edited by David Crystal. Cambridge UK: Press Syndicate of the University of Cambridge, 1992.
16. Carbonell, Jaime G. "Interlingua," Machine Translation Summit. 30-31. Tokyo: Ohmsha, Ltd., 1989.
17. Castellano, Lanna et al. "Session 5: Summary of discussion," Tools for the Trade, Translating and the Computer 5. 237-240. London: Aslib, 1985.
18. Childs, Larry. "The Ideal Machine Translation Environment," Proceedings of the 37th International Technical Communication Conference. RT2-RT4. Washington DC: Society for Technical Communication, 1990.
19. "Computer-Aided Translations," Information and Software Technology, 29: 51 (January/February 1987).
20. "Computer Translation by Telephone," Electronics and Wireless World, 93: 1071 (October 1987).
21. "Conversational Computers," IBM Directions, 4: 22-25 (Fall 1990).
22. "Current Machine Translation Systems in Japan," Machine Translation Summit. 213-224. Tokyo: Ohmsha, Ltd., 1989.
23. Cutler, Blayne. "Parlez-Vous Fujitsu?," American Demographics, 10: 56 (September 1988).
24. Dambrot, Stuart M. "Translating Systems in Japan," Datamation, 35: 102-13 - 102-14 (1 November 1989).
25. Dorr, Bonnie Jean. INITRAN: An Interlingual Machine Translation System. A.I. Memo No. 998. Contract N00014-80-C-0505. Cambridge MA: Massachusetts Institute of Technology Artificial Intelligence Laboratory, December 1987 (AD-A193 631).
26. "Early Natural Language Processing Successes," Computing Canada, 17: 25 (10 October 1991).
27. "Electronic Phrase Book from British Telecom," Electronics and Power, 33: 540 (September 1987).
28. Etheridge, James. "Eurotra: Speeding Toward 1992," Datamation, 35: 102-5 - 102-6 (1 November 1989).

29. Ferkinhoff, Thomas and David Vaughan. "Readability Formulas and Technical Order Preparation in the United States Air Force," Proceedings of the 37th International Technical Communication Conference. RT116-RT118. Washington DC: Society for Technical Communication, 1990.
30. Fowler, H.W. A Dictionary of Modern English Usage (Second Edition). Oxford UK: Oxford University Press, 1985.
31. Fuchi, Kazuhiro. "The FGCS Project and Machine Translation Systems," Machine Translation Summit. 13-20. Tokyo: Ohmsha, Ltd., 1989.
32. Gardner, Dana. "Machines Break the Language Barrier," Design News, 45: 22-23 (27 March 1989).
33. Gebruers, Rudi. "Valency and MT: Recent Developments in the METAL System," Second Conference on Applied Natural Language Processing. 168-175. Morristown NJ: Association for Computational Linguistics, 1988.
34. Gingras, Becky. "Simplified English in Technical Manuals," Technical Communication, 34: 24-28 (1987).
35. Globalink Translation System, The. User's Guide and Reference Manual. Globalink, Inc., Fairfax VA, 1992.
36. Gouin, Capt Andre R. CAF. French to English Machine Translation System Based Upon Digital Computer Software Programs (Systran). MS Thesis AFIT/GE/BE/70-9. School of Engineering, Air Force Institute of Technology (AU), Wright-Patterson AFB OH, March 1970 (AD-874523).
37. GTS Communicating with the World. Advertising Brochure. Globalink, Inc., Fairfax VA, undated.
38. Hawes, Ralph E. "Logos: The Intelligent Translation System," Tools for the Trade, Translating and the Computer 5. 133-139. London: Aslib, 1985.
39. Henke, Kristine A. "Facilitating Translation of Technical Information for the International Market," Proceedings 2 1990 IEEE International Professional Communication Conference. 90-94. IEEE Press, 1990.
40. Horning, Beth. "Language Busters," Technology Review: 51-57 (October 1991).
41. Hussey, Tim and Mark Homnack. "Foreign Language Software Localization," Proceedings of the 37th International Technical Communication Conference. RT44-RT47. Washington DC: Society for Technical Communication, 1990.



42. Hutchins W. John. "Prospects In Machine Translation," Machine Translation Summit. 7-12. Tokyo: Ohmsha, Ltd., 1989.
43. Ikehara, Satoru. "Approach to Machine Translation Method Based on Constructive Process Theory," Review of the Electrical Communication Laboratories (Tokyo), 37: 39-44 (1 January 1989).
44. Inaba, Minoru. "Fujitsu Develops Japanese-English Machine Translator," Electronic News, 35: 19 (2 January 1989).
45. "Interactive Author-Assistance Tool," IBM Technical Disclosure Bulletin, 28: 5284-5286 (12 May 1986).
46. Isabelle, Pierre and Laurent Bourbeau. "TAUM-AVIATION: Its Technical Features and Some Experimental Results," Computational Linguistics, 11: 18-27 (January-March 1985).
47. Ishizaki, Shun. "Machine Translation Projects," Machine Translation Summit. 35-36. Tokyo: Ohmsha, Ltd., 1989.
48. Johnson, Tim. "NLP Takes Off," Datamation, 3: 91-93 (15 January 1986).
49. Johnstone, B. "The Language of Chip Making," Far Eastern Economic Review, 148: 70 (3 May 1990).
50. Kaji, Hiroyuki. "HICATS/JE: A Japanese-to-English Machine Translation System Based on Semantics," Machine Translation Summit. 101-106. Tokyo: Ohmsha, Ltd., 1989.
51. Kaye, Lincoln. "Babel In, Babel Out," Far Eastern Economic Review, 146: 95 (5 October 1989).
52. Kinkaid, Peter J. et al. "Computer Readability Editing System," IEEE Transactions on Professional Communications, 24: 38 (March 1981).
53. Kisik, Lee. "Korea," Machine Translation Summit. 172-174. Tokyo: Ohmsha, Ltd., 1989.
54. Kittredge, Richard I. "The Significance of Sublanguage for Automatic Translation," Proceedings of the Conference on Theoretical and Methodological Issues in Machine Translation of Natural Languages. 154-166. Hamilton NY: Colgate University, 1985.
55. Kohl, John R. "Using 'Syntactic Clues' to Enhance Readability for Non-Native Speakers of English," Proceedings of the International Technical Communication Conference. RT54-RT57. 1991.

56. Kubota, Akira. "Japan," Machine Translation Summit. 170-171. Tokyo: Ohmsha, Ltd., 1989.
57. Kurita, Shohei. "AI in Japan: Closing the Gap with the U.S.," Electronic Business, 14: 110-111 (15 January 1988).
58. Kusuda, Tamami. "United States of America," Machine Translation Summit. 177-180. Tokyo: Ohmsha, Ltd., 1989.
59. Kuzela, Lad. "Bilingual Machines," Industry Week, 237: 82-83 (17 October 1983).
60. Kwasny, Stan C. "A PDP Approach to Deterministic Natural Language Parsing," Neural Networks, 1: 305 (1988).
61. Lang, Francois-Michel. "PUNDIT Apprend le Francais: The PRATTFALL Machine-Translation Module," Fifth Annual AI Systems in Government Conference. 157-166. IEEE Press, 1990.
62. Lavroff, O. "Comparison Between Humans and Automatic Translations," AGARD Lecture Series No. 171, Benefits of Computer Assisted Translation to Information Managers and End-Users. 10-9-10-38. Neuilly sur Seine France: AGARD, 1990 (AD-A225 965).
63. Lufkin, James M. "Current Trends in Technical Translation," 1989 IEEE International Professional Communication Conference. 238-243. IEEE Press, 1989.
64. Lytinen, Steven L. "Integrating Syntax and Semantics," Proceedings of the Conference on Theoretical and Methodological Issues in Machine Translation of Natural Languages. 167-178. Hamilton NY: Colgate University, 1985.
65. Madlin, Nancy. "The New Computer Lingo: Foreign Language Translations," Management Review, 77: 58-60 (July 1988).
66. Magnusson-Murray, Ulla. "Operational Experience of a Machine Translation Service," Tools for the Trade, Translating and the Computer 5. 171-180. London: Aslib, 1985.
67. McClave, James T. and P. George Benson. Statistics for Business and Economics (Fifth Edition). San Francisco: Dellen Publishing Company, 1991
68. Mendosa, Rick. "Translator's Little Helpers," Hispanic Business: 32-34 (October 1991).

69. Meyer, Richard. "Word Crunching," Financial World, 159: 76-77 (4 July 1990).
70. Moos, W.H.I., Program Manager, Military Programs. Telephone interview. Globalink Inc., Fairfax VA, 24 August 1992.
71. Muraki, Kazanori. "PIVOT: Two-Phase Machine Translation System," Machine Translation Summit. 113-115. Tokyo: Ohmsha, Ltd., 1989.
72. ----- . "VENUS: Two-Phase Machine Translation System," Future Generations Computer Systems, 2: 117-119 (1986).
73. Nagao, Makato. "Current Status and Future Trends in Machine Translation," Future Generations Computer Systems, 2: 78-82 (June 1986).
74. ----- et al. "Dealing with Incompleteness of Linguistic Knowledge in Language Translation," Proceedings of COLING 84. 420-427 (1984).
75. ----- et al. "Japanese Government Project for Machine Translation, The," Journal of the Association of Computational Linguistics, 11: 91-110 (1985).
76. ----- et al. "Machine Translation from Japanese into English," Proceedings of the IEEE, 74: 993-1012 (July 1986).
77. ----- . "Progress of Artificial Intelligence in Japan," International Workshop on Artificial Intelligence for Industrial Applications. 3-6. IEEE Press, 1988.
78. ----- . "Report on Machine Translation Summit," Machine Translation Summit. 181-186. Tokyo: Ohmsha, Ltd., 1989.
79. ----- et al. "Science and Technology Agency's Mu Machine Translation Project," Future Generations Computer Systems, 2: 125-139 (1986).
80. Nakamura, Jun-ichi. et al. "GRADE: A Software Environment for Machine Translation," Computers and Translation, 3: 69-82 (1988).
81. Nash, Jim. "The Many Tongues of Computers," Computerworld, 25: 20 (18 February 1991).
82. Newquist, H.P. "Breaking a Barrier to Trade," Computerworld, 26: 33 (24 February 1991).
83. Nitta, Yoshihiko. "Problems of Machine Translation Systems: Effect of Cultural Differences on Sentence Structure," Future Generations Computer Systems, 2: 101-115 (June 1986).
84. Nomura, Hiroshiro. "Technical Prospects of Machine Translation," Machine Translation Summit. 21-24. Tokyo: Ohmsha, Ltd., 1989.

85. Noordzij, Jaap. "Blackboard Translation," Systems International, 15: 57-60 (March 1987).
86. "Now Cheerful Inhuman Word Swap a Wonder Be," The Economist, 305: 95-96 (14 November 1987)
87. Oshio, Tadaaki. "Applied Testing of HICATS/JE for Japanese Patent Abstracts," Machine Translation Summit. 140-144. Tokyo: Ohmsha, Ltd., 1989.
88. Perschke, Sergei. "EUROTRA Project," Machine Translation Summit. 37-44. Tokyo: Ohmsha, Ltd., 1989.
89. Peterson, D.A.T. "Approaches to Computer-Assisted Translation in Canada," Technical Communication, 35: 35-36 (First Quarter 1988).
90. Piggot, Ian M. "Recent Developments in Practical Machine Translation," Tools for the Trade, Translating and the Computer 5. 97-103. London: Aslib, 1985.
91. ----- "Typology of Existing Systems," AGARD Lecture Series No. 171, Benefits of Computer Assisted Translation to Information Managers and End-Users. 1-1-1-5. Neuilly sur Seine France: AGARD, 1990 (AD-A225 965).
92. Pinna, D.J. "Problems Encountered with the Use of Computer Assisted Translation in a Technical Publications Production Environment," AGARD Lecture Series No. 171, Benefits of Computer Assisted Translation to Information Managers and End-Users. 4-1-4-9. Neuilly sur Seine France: AGARD, 1990 (AD-A225 965).
93. Pons, Henrietta B. "WCC's Translation Service Bureau," Tools for the Trade, Translating and the Computer 5. 165-169. London: Aslib, 1985.
94. Raman, S. and N. Alwar. "An AI-Based Approach to Machine Translation in Indian Languages," Communications of the ACM, 33: 521-527 (May 1990).
95. Redish, Janice C. and Jack Selzer. "The Place of Readability Formulas in Technical Communication," Technical Communication, 32: 46-52 (1985).
96. Rolling, Loll. "Concluding Remarks 1," Machine Translation Summit. 187-191. Tokyo: Ohmsha, Ltd., 1989.
97. ----- "European Community," Machine Translation Summit. 164-165. Tokyo: Ohmsha, Ltd., 1989.

98. Rossetto, Louis. "The Next Computer Revolution: Wordcrunching," The Electronic Library, 7: 301-304 (October 1989).
99. Ryan, Joann P. "SYSTRAN: A Machine Translation System to Meet User Needs," Machine Translation Summit. 116-121. Tokyo: Ohmsha, Ltd., 1989.
100. Sager, J.C. "Chairman's Introduction: Naming Machine Translation," Tools for the Trade, Translating and the Computer 5. 93-95. London: Aslib, 1985.
101. Schneider, Thomas. "End Users' Needs," AGARD Lecture Series No. 171, Benefits of Computer Assisted Translation to Information Managers and End-Users. 5-1-5-4. Neuilly sur Seine France: AGARD, 1990 (AD-A225 965).
102. -----, "Metal System, Status 1987, The," Machine Translation Summit. 122-127. Tokyo: Ohmsha, Ltd., 1989.
103. -----, "State of the Art in West Germany," AGARD Lecture Series No. 171, Benefits of Computer Assisted Translation to Information Managers and End-Users. 8-1-8-12. Neuilly sur Seine France: AGARD, 1990 (AD-A225 965).
104. Shah, Raymond. "Translation of Engineering Documentation with METAL," Machine Translation Summit. 152-159. Tokyo: Ohmsha, Ltd., 1989.
105. Simmons, R.F. "Technologies for Machine Translation," Future Generations Computer Systems, 2: 83-94 (June 1986).
106. Slocum, Jonathon. "A Survey of Machine Translation: Its History, Current Status, and Future Prospects," Computational Linguistics, 11: 1-17 (January-March 1985).
107. "Software Improves Foreign Translations," Machine Design, 62: 147 (11 January 1990).
108. Sophonpanich, K. "Thailand," Machine Translation Summit. 175-176. Tokyo: Ohmsha, Ltd., 1989.
109. Spencer, Kathy and Peggy Yates. "Tell it to the World: Technical Communication in the Global Marketplace," Proceedings of the 37th International Technical Communication Conference. RT149-RT152. Washington DC: Society for Technical Communication, 1990.
110. Statistix, User's Manual. Edited by Joan Siegel. Analytical Software, St. Paul MN, 1992.

111. Stone, Richard. "The Education of Silicon Linguists," Science, 253: 854-855 (23 August 1991).
112. Streiff, A.A. "New Developments in TITUS 4," Tools for the Trade, Translating and the Computer 5. 185-192. London: Aslib, 1985.
113. Sudarwo, Iman. "Indonesia," Machine Translation Summit. 168. Tokyo: Ohmsha, Ltd., 1989.
114. Sutherland, Allison, Linguist. Personal Notes. Globalink, Inc., Fairfax VA, undated.
115. Tanaka, Hozumi. "Commercial Machine Translation Systems," Machine Translation Summit. 91-92. Tokyo: Ohmsha, Ltd., 1989.
116. Tarantola, Carlo A. "Machine Translation Past, Present, and Future," IEEE Expert Magazine, 3: 91 (Spring 1988).
117. Tenney, Merle D. "Machine Translation, Machine-Aided Translation, and Machine-Impeded Translation," Tools for the Trade, Translating and the Computer 5. 105-113. London: Aslib, 1985.
118. Thwaite, David. "Automatic Interpreters," Systems International, 16: 47-48 (February 1988).
119. -----, "Whipping Document Deadlines with CAT," Systems International, 16: 33-34 (March 1988).
120. Tomita, Masaru and Jaime G. Carbonell. "CMU Project," Machine Translation Summit. 72-81. Tokyo: Ohmsha, Ltd., 1989.
121. -----, "Sentence Disambiguation by Asking," Computers and Translation, 1: 39-51 (January-March 1986).
122. Tschira, Klaus E. "Looking Back at a Year of German-English MT with Logos," Tools for the Trade, Translating and the Computer 5. 215-225. London: Aslib, 1985.
123. Tsui, Ming. "Canada," Machine Translation Summit. 163. Tokyo: Ohmsha, Ltd., 1989.
124. Tsuji, Yoshihide. "Machine Translation with Japan's Neighboring Countries," Machine Translation Summit. 50-53. Tokyo: Ohmsha, Ltd., 1989.
125. Uchida, Hiroshi. "ATLAS: Automatic Translation System," Fujitsu Scientific and Technical Journal, 21: 317-329 (July 1985).

126. -----. "ATLAS II: A Machine Translation System Using Conceptual Structure as an Interlingua," Machine Translation Summit. 93-100. Tokyo: Ohmsha, Ltd., 1989.
127. -----. "Fujitsu Machine Translation System: ATLAS," Future Generations Computer Systems, 2: 95-100 (1986).
128. "Update of Machine Translation of Natural Languages," Computers and People, 35: 24-26 (May-June 1986).
129. Varile, Giovanni B. and Peter Lau. "Practical Experience With a Multilingual Machine Translation System Under Development," Second Conference on Applied Natural Language Processing. 160-167. Morristown NJ: Association for Computational Linguistics, 1988.
130. Vasconcellos, Muriel. "Management of the Machine Translation Environment: Interaction of Functions at the Pan American Health Organization," Tools for the Trade, Translating and the Computer 5. 115-129. London: Aslib, 1985.
131. Vauquois, Bernard and Christian Boitet. "Automated Translation at Grenoble University," Computational Linguistics, 11: 28-36 (January-March 1985).
132. Velte, Charles. "Is Bad Breadth Preventing Effective Translation of Your Text?," Proceedings of the 37th International Technical Communication Conference. RT5-RT7. Washington DC: Society for Technical Communication, 1990.
133. Wagner, Elizabeth. "Rapid Post-Editing of Systran," Tools for the Trade, Translating and the Computer 5. 199-214. London: Aslib, 1985.
134. Warotamasikkhadit, Udom. "Computer Aided Translation Project, University Sains Malaysia, Penang, Malaysia," Computers and Translation, 1: 113 (April-June 1986).
135. Whenmouth, Edwin. "Interpreting, Japanese Style," Industry Week, 238: 53-54 (5 June 1989).
136. Wilson, John. "Translating Speech by Computer," Electronics and Wireless World, 93: 1100 (November 1987).
137. Yamaoka, Yoichi. "Computer Aided Translation System and Computerized Dictionary," Machine Translation Summit. 145-146. Tokyo: Ohmsha, Ltd., 1989.
138. Zaretsky, Gregory. "What's New in Machine Translation," Communication World, 7: 31-34 (August 1990).

### Vita

Squadron Leader Ernie Walsh was born on 19 September 1960 in Brisbane, the state capital of Queensland, Australia. He graduated from Kelvin Grove High School in 1977 and entered the Royal Australian Air Force Academy in January 1978. He graduated from the Academy in December 1982 with a Bachelor of Science in Physics, a Graduate Diploma in Military Aviation, and a permanent commission. His first tour was spent at Maintenance Squadron East Sale, where he filled a number of maintenance oriented positions until being posted to Air Force Office in January of 1986. During his three years at Air Force Office he worked in the Directorate of Maintenance Policy and was involved in various acquisition projects including the PC9/A trainer aircraft. He was also responsible for the Engineering Division computer systems and introduced extensive unix and PC based office support during his tenure. His next posting took him to Perth, Western Australia in January 1989, where he served as the RAAF Resident Engineer. Here, he was responsible for liaising with Perth area maintenance and engineering contractors in support of RAAF and Royal Australian Navy maintenance contracts. He entered the Information Resource Management program at the School of Systems and Logistics, US Air Force Institute of Technology, in May 1991.

Permanent Address:	10 Siesta Street CAMIRA QLD 4300 Australia
--------------------	--



# REPORT DOCUMENTATION PAGE

Form Approved  
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)		2. REPORT DATE December 1992	3. REPORT TYPE AND DATES COVERED Master's Thesis	
4. TITLE AND SUBTITLE IMPROVING THE QUALITY OF MACHINE TRANSLATIONS BY PROVIDING STYLISTIC GUIDELINES TO AUTHORS AND EDITORS			5. FUNDING NUMBERS	
6. AUTHOR(S) Edward J. Walsh, Squadron Leader, Royal Australian Air Force				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Air Force Institute of Technology, WBAFB OH 45433-6583			8. PERFORMING ORGANIZATION REPORT NUMBER AFIT/GIR/DEM/92D-9	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)			10. SPONSORING / MONITORING AGENCY REPORT NUMBER	
11. SUPPLEMENTARY NOTES				
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution unlimited			12b. DISTRIBUTION CODE	
13. ABSTRACT (Maximum 200 words) This research sought to identify stylistic guidelines which could be used by the authors and editors of English language texts in order to improve the quality of subsequent machine translations to other languages. A literature review was conducted to provide background information and to determine what guidelines would be appropriate. Nine guidelines were developed on the basis of the review. The guidelines were then tested. They were applied to five paragraphs from a Royal Australian Air Force technical publication by the researcher and an independent volunteer. Both the resulting texts and the original paragraphs were translated into French using the Globalink Translation System and the quality of the translations obtained was evaluated by four human translators. While the translations of the paragraphs which had been edited by the volunteer were significantly better than the translations of the original texts, the translations of the paragraphs edited by the researcher were not. Use of the guidelines by authors and editors cannot be recommended on the basis of this result. However, analysis of the conduct and results of the research suggested potentially fruitful alternative approaches to the development and testing of guidelines. These approaches are presented as specific recommendations for further research.				
14. SUBJECT TERMS Computational Linguistics, Documents, Language Translation, Machine Translation, Natural Language, Text Processing			15. NUMBER OF PAGES 154	
			16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT UL	

## AFTT RESEARCH ASSESSMENT

The purpose of this questionnaire is to determine the potential for current and future applications of AFTT thesis research. Please return completed questionnaires to: AFTT/LSC, Wright-Patterson AFB OH 45433-9905.

1. Did this research contribute to a current research project?

a. Yes

b. No

2. Do you believe this research topic is significant enough that it would have been researched (or contracted) by your organization or another agency if AFTT had not researched it?

a. Yes

b. No

3. The benefits of AFTT research can often be expressed by the equivalent value that your agency received by virtue of AFTT performing the research. Please estimate what this research would have cost in terms of manpower and/or dollars if it had been accomplished under contract or if it had been done in-house.

Man Years \_\_\_\_\_ \$ \_\_\_\_\_

4. Often it is not possible to attach equivalent dollar values to research, although the results of the research may, in fact, be important. Whether or not you were able to establish an equivalent value for this research (3, above) what is your estimate of its significance?

a. Highly  
Significant

b. Significant

c. Slightly  
Significant

d. Of No  
Significance

5. Comments

\_\_\_\_\_  
Name and Grade

\_\_\_\_\_  
Organization

\_\_\_\_\_  
Position or Title

\_\_\_\_\_  
Address